

STIC-Biotech/ChemLib

94181

From: Jiang, Dong  
Sent: Thursday, May 15, 2003 6:49 PM  
To: STIC-Biotech/ChemLib  
Subject: 10/050,704

RECEIVED  
Priority: 4/9/99  
MAY 16 2003  
HGS (Soppet, D.)  
(STIC)

Please search SEQ ID NO:125

-issued & Pub.  
-commercial

no utility  
Point of Contact:  
Toby Port  
Technical Info. Specialist  
CM1 6A04  
703-305-3534  
# - 1/18/02

Please send results on paper to Dong Jiang in 10D-08 (mail stop CM1-10D19).  
Thank you very much.

Dong Jiang (78243)  
703-305-1345  
U.S. Patent and Trademark Office  
Art Unit 1646  
dong.jiang@uspto.gov  
CM1-10D08  
Mail stop: CM1-10D19

Do a alignment b/w SEQ 125 & 127.

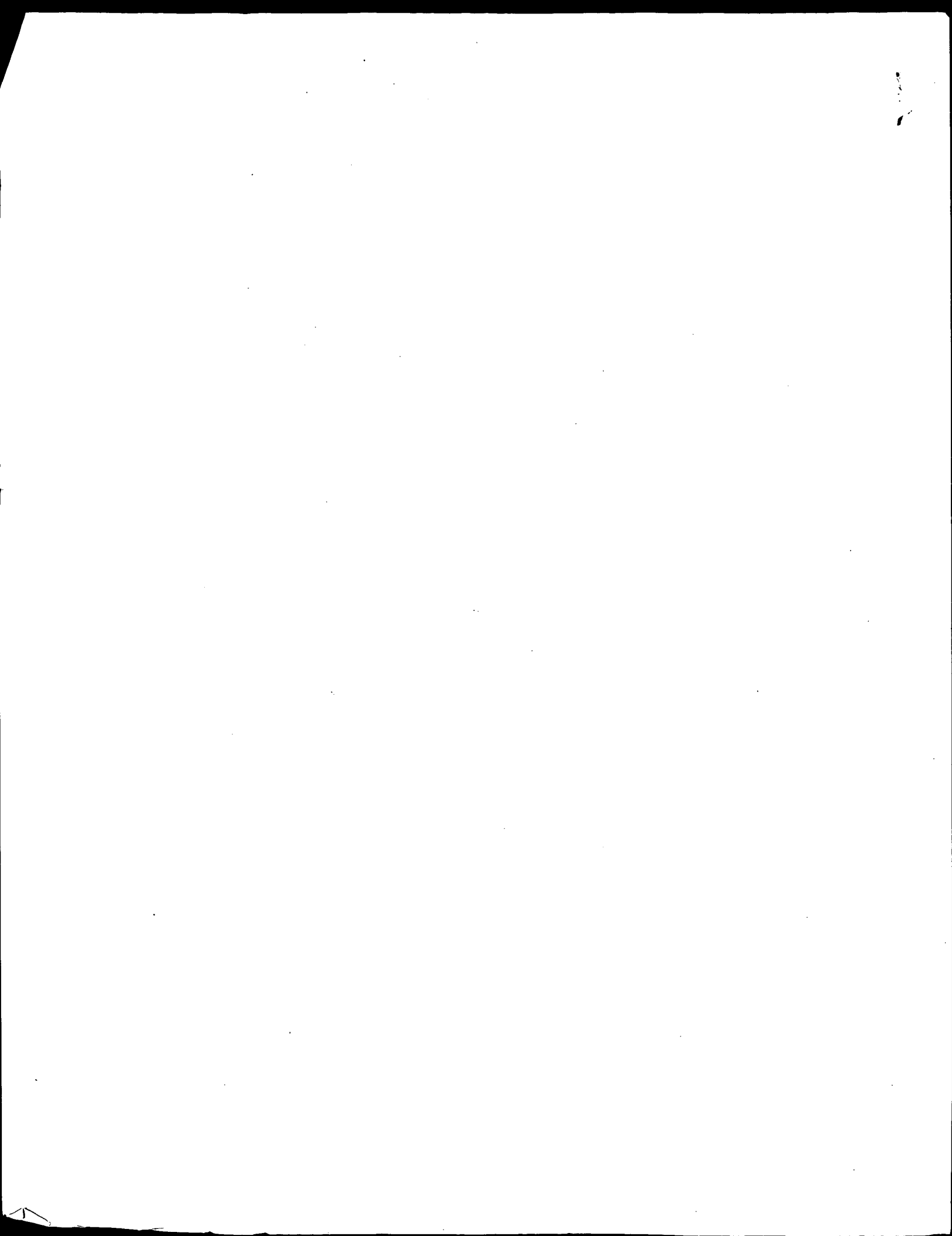
Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 5/16  
Date Completed: 5/16  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:

NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: May 16, 2003, 10:45:57 ; Search time 40 Seconds  
(without alignments)  
2088.703 Million cell updates/sec

Title: US-10-050-704-125

Perfect score: 3208

Sequence: 1 MEARVHALKRGVSLICVF.....KATEPGLFGVMAHNGEVRTA 627

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A.GeneSeq.101002.\*  
1: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3208	100.0	627	21	Human secreted pro
2	2766	86.2	889	22	Novel human diago
3	2587.5	80.7	834	22	Novel human diago
4	2473.5	77.1	653	22	Novel human diago
5	2304	71.8	797	22	Novel human diago
6	2287	71.3	787	17	Human protocadheri
7	2249	70.1	800	21	Human protocadheri
8	2249	70.1	800	22	Human protocadheri
9	2237	69.7	1888	22	Novel human diago
10	2142	66.8	798	20	Human PRO531 prote

11	2142	66.8	798	21	AA844302
12	2142	66.8	798	21	AA824049
13	2142	66.8	798	21	AA833433
14	2142	66.8	798	23	AA838657
15	2118	66.0	779	21	AA849997
16	2009.5	62.6	797	17	AA868666
17	1879	58.6	744	22	AA894986
18	1841	57.4	743	21	AA891492
19	1576.5	49.1	592	22	AA894487
20	1558	48.6	554	23	AB842004
21	1457	45.4	559	23	AB872287
22	1418	44.2	931	22	AA878649
23	1418	44.2	949	22	AA812315
24	1418	44.2	949	22	AA879633
25	1401	43.7	833	22	AB826040
26	1400	43.6	858	22	AA879446
27	1393	43.4	931	22	AA878462
28	1383.5	43.1	932	22	AB817156
29	1377.5	42.9	832	22	AB823820
30	1343.5	41.9	916	20	AA841743
31	1343.5	41.9	916	21	AA844299
32	1343.5	41.9	916	22	AA829062
33	1323	41.2	820	22	AB823824
34	1288	40.1	1958	22	AB821821
35	1268	39.5	459	22	AA892859
36	1257	39.2	405	22	AA893330
37	1232.5	38.4	934	23	AA847925
38	1217.5	38.0	1008	22	AB815334
39	1217	37.9	682	15	AA849144
40	1217	37.9	682	17	AA887154
41	1217	37.9	836	15	AA858912
42	1217	37.9	836	17	AA887153
43	1217	37.9	904	15	AA858907
44	1217	37.9	904	17	AA887147
45	1192.5	37.2	949	22	AB823232

## ALIGNMENTS

RESULT 1	AA838349	standard; Protein; 627 AA.
XX	AA838349:	
XX	31-JAN-2001	(first entry)
DE	Human secreted protein encoded by gene 29 clone HOFND85.	
XX	Immunosuppressive; antiarthritic; antirheumatic; antiproliferative;	
KW	cytostatic; cardiant; vasotropic; cerebroprotective; neuroprotective;	
KW	neotropic; antibacterial; virucide; fungicide; ophthalmological; human;	
KW	vulnary; gene therapy; infection; secreted protein.	
OS	Homo sapiens.	
XX	MO200061623-A1.	
PN	19-OCT-2000.	
PD	06-APR-2000; 2000WO-US08979.	
XX	09-APR-1999; 99US-0128693.	
PR	26-APR-1999; 99US-0130991.	
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	Ruben SM, Ni J, Komatsu GA, Soppet DR, Shi Y;	
PI	Latif DM, Olsen HS, Ebner R, Florence KA, Moore PA, Bltse CE;	
PI	Young FE;	
DR	WPI, 2000-647418/62.	

XX New nucleic acid molecules encoding 62 human secreted proteins for  
PT diagnosing, preventing, treating or ameliorating medical conditions and  
PT used as food additives or preservatives

PS Claim 11; Page 619-621; 716pp; English.

XX Sequences AAB38321-B38396 represent the amino acid sequences of 62  
CC human secreted proteins encoded by the genes AAG69512-C69587. The genes  
CC and proteins are useful for preventing, ameliorating or treating medical  
CC conditions, e.g. by protein or gene therapy. The genes are isolated from  
CC a range of human tissues disclosed in the specification. The nucleic  
CC acids, proteins, antibodies and (ant)agonists are useful in the  
CC diagnosis, treatment and prevention of: (a) autoimmune diseases e.g.  
CC rheumatoid arthritis; (b) hyperproliferative disorders e.g. neoplasms  
CC of the breast or liver; (c) cardiovascular disorders e.g. cardiac  
CC arrest; (d) cerebrovascular disorders e.g. cerebral ischemia; (e)  
CC angiogenesis; (f) nervous system disorders e.g. Alzheimer's disease; (g)  
CC infections caused by bacteria, viruses and fungi; and (h) ocular  
CC disorders e.g. corneal infection. The polypeptides can also be used to  
CC aid wound healing and epithelial cell proliferation, to prevent skin  
CC aging due to sunburn, to maintain organs before transplantation, for  
CC supporting cell culture of primary tissues, to regenerate tissues and in  
CC chemotaxis.

XX Sequence 627 AA:

Query Match 100.0%; Score 3208; DB 21; Length 627;  
Best Local Similarity 100.0%; Pred. No. 6.5e-278;  
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEARVHALOKROYSLVCLVGVSMAGAEPLRFVAEETGFFLANLAIDGLGVEELS 60  
DB 1 MEARVHALOKROYSLVCLVGVSMAGAEPLRFVAEETGFFLANLAIDGLGVEELS 60  
QY 61 ARGRIVSDEITGFLNPLTGLDLNENKLDRELCPPEPCVLPOLLLEPFQFRAE 120  
DB 61 ARGRIVSDEITGFLNPLTGLDLNENKLDRELCPPEPCVLPOLLLEPFQFRAE 120  
QY 121 LWRDINDHSVFLDREITNLIESTPGATFLLSADSDVGINLRYTSSNVYFHI 180  
DB 121 LWRDINDHSVFLDREITNLIESTPGATFLLSADSDVGINLRYTSSNVYFHI 180  
QY 181 NVHDNEGNYSELVLDKVDREVEBELRLTGLDGGSPSSGTLRLIIVLDINDVP 240  
DB 181 NVHDNEGNYSELVLDKVDREVEBELRLTGLDGGSPSSGTLRLIIVLDINDVP 240  
QY 241 EFVESLYKQVPEPNSPVGSLVYVSARDIDTGSNGEIVAFYATERTLKTFRINSTGN 300  
DB 241 EFVESLYKQVPEPNSPVGSLVYVSARDIDTGSNGEIVAFYATERTLKTFRINSTGN 300  
QY 301 LHKALNEALQTYTLTQAKDGGGLSKCTVVVAVHTDINDPPELMSLTSPSPENS 360  
DB 301 LHKALNEALQTYTLTQAKDGGGLSKCTVVVAVHTDINDPPELMSLTSPSPENS 360  
QY 361 PETVVAVFRIRDRSGNNKMYCSIODHLPEVLKPSVEVEFTLVTERALDREEREYNT 420  
DB 361 PETVVAVFRIRDRSGNNKMYCSIODHLPEVLKPSVEVEFTLVTERALDREEREYNT 420  
QY 421 IVTDTGLRKLQHNLTJVSVDVNDNAFTSQTTYTLKVRNNSPALHIGSVATDSDS 480  
DB 421 IVTDTGLRKLQHNLTJVSVDVNDNAFTSQTTYTLKVRNNSPALHIGSVATDSDS 480  
QY 481 GANAQVTSGLRPHDQPLGSLVSIINADNGOLFALRSIDFALQAFERRGADRGSPA 540  
DB 481 GANAQVTSGLRPHDQPLGSLVSIINADNGOLFALRSIDFALQAFERRGADRGSPA 540  
QY 541 LSSQALVRVLVADANDNAFVLYPLONGSAPCTELVPRAAEGYVAVVAVDGGSGONA 600  
DB 541 LSSQALVRVLVADANDNAFVLYPLONGSAPCTELVPRAAEGYVAVVAVDGGSGONA 600  
QY 601 WLSTYOLKATPEGLFGVMAHNGEVRTA 627  
DB 601 WLSTYOLKATPEGLFGVMAHNGEVRTA 627

DB 601 WLSTYOLKATPEGLFGVMAHNGEVRTA 627

RESULT 2

ID ABG23879  
ABG23879 standard; Protein; 889 AA.

XX ABG23879;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #23870.

XX Human: chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HXSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS88066.

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity

PT Claim 20; SEQ ID NO 54238; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX and gene mapping, and in recombinant production of (II). The  
XX polynucleotides are also used in diagnostics as expressed sequence tags  
XX for identifying expressed genes. (I) is useful in gene therapy techniques  
XX to restore normal activity of (II) or to treat disease states involving  
XX (II). (II) is useful for generating antibodies against it, detecting or  
XX quantitating a polypeptide in tissue, as molecular weight markers and as  
XX a food supplement. (II) and its binding partners are useful in medical  
XX imaging of sites expressing (II). (I) and (II) are useful for treating  
XX disorders involving aberrant protein expression or biological activity.  
XX The polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. ABG00010-ABG30377 represent novel human  
XX diagnostic amino acid data for the invention.  
XX Note: The sequence data for this patent did not appear in the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 889 AA:

Query Match 86.2%; Score 2766; DB 22; Length 889;  
Best Local Similarity 85.8%; Pred. No. 4.2e-238;

Matches 538; Conservative 32; Mismatches 57; Indels 0; Gaps 0;

QY 1 MEARVHALOKROYSLVCLVGVSMAGAEPLRFVAEETGFFLANLAIDGLGVEELS 60  
DB 95 MEARVERAVQKQVFLVCLVGVSMAGAEPLRFVAEETGFFLANLAIDGLGVEELS 154

QY 61 ARGCRIVSDETIGELLNPLTGDLINKRELCGPTPECVLPFCULLKPKQIFRAE 120  
 DB 155 AGSTRIVSDQNMQLLSLSTGDLINKRELCGPREPCVLPFCULLKPKQIFRAE 214  
 QY 121 LMRVINDHSPVFLDREITLNLSTPGATFLLESANDSDVGINNLNNTISSNVEYFI 180  
 DB 215 LMRVINDHSPVFLDREITLNLSTPGATFLLESANDSDVGINNLNNTISSNVEYFI 274  
 QY 181 NVHNDGEGNVSELYDKVLDREYPELRLTLTGDDGSPRSGTTLRLIYLDINDVNP 240  
 DB 275 NVHNDGEGNVSELYDKVLDREYPELRLTLTGDDGSPRSGTTLRLIYLDINDVNP 334  
 QY 241 EVESTLYKVQVPENSPVSGSVVTSARLDITGSGEIVYAFYATERTLKTFRINSTSGN 300  
 DB 335 DEVRSLTKYQVPENSPVSGSVVTSARLDITGSGEIVYAFYATERTLKTFRINSTSGN 394  
 QY 301 LHLKAELEAYEAIQTYTLTIOAKDGGGLSGCTVYVAVTDINDNPELLMSLTSPIPENS 360  
 DB 395 LHLKAELEAYEAIQTYTLTIOAKDGGGLSGCTVYVAVTDINDNPELLMSLTSPIPENS 454  
 QY 361 PETVAVAFRIIRDSDGNNAKMWCSIODHLPEYKLSVENEFTLVTERALDREEREYNT 420  
 DB 455 PETVAVAFRIIRDSDGNNAKMWCSIODHLPEYKLSVENEFTLVTERALDREEREYNT 514  
 QY 421 ITVTDGTRPLKTOHNLTVTSVDNDNAPTFESQTYTLTVRENNSPALHIGSVSATDRDS 480  
 DB 515 ITVTDGTRPLKTOHNLTVTSVDNDNAPTFESQTYTLTVRENNSPALHIGSVSATDRDS 574  
 QY 481 GANAQYITSLPRLPHDPLGLSVISINADNGOLFALRSIDFEALQAFERRVGAADGSPA 540  
 DB 575 GINAQYITSLPRLPHDPLGLSVISINADNGOLFALRSIDFEALQAFERRVGAADGSPA 634  
 QY 541 LSSQALVRYLVANANAPFVLPLONGSAPCELPRAAEGYLVAKVVAADGDSGQNA 600  
 DB 635 LSSQALVRYLVANANAPFVLPLONGSAPCELPRAAEGYLVAKVVAADGDSGQNA 694  
 QY 601 WLSTYQLLKATEPGLFGVMAHNGEVRTA 627  
 DB 695 WLSTYQLLKATEPGLFGVMAHNGEVRTA 721  
 RESULT 3  
 ID ABG23876 standard: Protein; 834 AA.  
 AC ABG23876:  
 DT 18-FEB-2002 (first entry)  
 XX Novel human diagnostic protein #23867.  
 DE Human: chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX Homo sapiens.  
 OS WO200175067-A2.  
 PN 11-OCT-2001.  
 PD 30-MAR-2001; 2001WO-US08631.  
 FE 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX (HSE-) HYSBO INC.  
 PA Dimaenac RT, Liu C, Tang YT;  
 PI WPI: 2001-639362/73.  
 DR N-PSDB: AAS88063.  
 XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity  
 XX Claim 20: SEQ ID No 54235; 103pp; English.  
 PS  
 XX The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pcl\_sequences.  
 XX  
 SO Sequence 834 AA:  
 Query Match 80.7%; Score 2587.5; DB 22; Length 834;  
 Best Local Similarity 74.9%; Pred. No. 3.6e-22;  
 Matches 517; Conservative 37; Mismatches 61; Indels 75; Gaps 3;  
 QY 1 MEARVHALOKROVSLICVFGLGVSWAGAEPLRYFAEETERTGFLANLADIGVEELS 60  
 DB 1 MEARVHALOKROVSLICVFGLGVSWAGAEPLRYFAEETERTGFLANLADIGVEELS 60  
 QY 61 ARGCRIVSDETIGELLNPLTGDLINKRELCGPTPECVLPFCULLKPKQIFRAE 120  
 DB 61 ARGCRIVSDQNMQLLSLSTGDLINKRELCGPREPCVLPFCULLKPKQIFRAE 120  
 QY 121 LMRVINDHSPVFLDREITLNLSTPGATFLLESANDSDVGINNLNNTISSNVEYFI 180  
 DB 121 LMRVINDHSPVFLDREITLNLSTPGATFLLESANDSDVGINNLNNTISSNVEYFI 180  
 QY 181 NVHNDGEGNVSELYDKVLDREYPELRLTLTGDDGSPRSGTTLRLIYLDINDVNP 240  
 DB 181 NVHNDGEGNVSELYDKVLDREYPELRLTLTGDDGSPRSGTTLRLIYLDINDVNP 240  
 QY 241 EVESTLYKVQVPENSPVSGSVVTSARLDITGSGEIVYAFYATERTLKTFRINSTSGN 300  
 DB 241 DEVRSLTKYQVPENSPVSGSVVTSARLDITGSGEIVYAFYATERTLKTFRINSTSGN 300  
 QY 301 LHLKAELEAYEAIQTYTLTIOAKDGGGLSGCTVYVAVTDINDNPELLMSLTSPIPENS 360  
 DB 301 LHLKAELEAYEAIQTYTLTIOAKDGGGLSGCTVYVAVTDINDNPELLMSLTSPIPENS 360  
 QY 361 PETVAVAFRIIRDSDGNNAKMWCSIODHLPEYKLSVENEFTLVTERALDREEREYNT 420  
 DB 361 PETVAVAFRIIRDSDGNNAKMWCSIODHLPEYKLSVENEFTLVTERALDREEREYNT 420  
 QY 421 ITVTDGTRPLKTOHNLTVTSVDNDNAPTFESQTYTLTVRENNSPALHIGSVSATDRDS 480  
 DB 421 ITVTDGTRPLKTOHNLTVTSVDNDNAPTFESQTYTLTVRENNSPALHIGSVSATDRDS 480  
 QY 481 GANA-----  
 DB 481 GANAQYITSLPRLPHDPLGLSVISINADNGOLFALRSIDFEALQAFERRVGAADGSPA 540  
 QY 485 -----QVTSLLPRLPHDPLGLSVISINADNGOLFALRSIDFEALQAFERRVGAADGSPA 536  
 DB 541 VSTAETNIGVTSLLPRLPHDPLGLSVISINADNGOLFALRSIDFEALQAFERRVGAADGSPA 600

QY 537 GSPALSSQALVRLVADANDNAPFLVPLQNGSAPCTELVPR-AAAGYLVAKVAVDGD 595  
 DB 601 GSTALSSQALVRLVADANDSSLFVLPPLQNGSAPCTELMRAAERGGLVTK----- 653  
 QY 596 SGNAMLSYQLLKATEPGLFGVMAHNGEVR 625  
 DB 654 ---NMLSSYQLLKATEPGLFGVMAHNGTDR 680

RESULT 4  
 ABG23877  
 ID ABG23877 standard; Protein: 653 AA.  
 AC ABG23877:  
 DT 18-FEB-2002 (first entry)  
 DE Novel human diagnostic protein #23868.  
 KW Human: chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 OS Homo sapiens.  
 PN WO200175067-A2.  
 PD 11-OCT-2001.  
 PF 30-MAR-2001; 2001WO-US08631.  
 PR 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 PA (HYSE-) HYSEQ INC.  
 PI Drmanac RT, Liu C, Tang YT;  
 DR WPI; 2001-639362/73.  
 DR N-PSDB; AAS68064.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -

Claim 20: SEQ ID No 54236; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations in  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 CC Sequence 653 AA:  
 Query Match 77.1%; Score 2473.5; DB 22; Length 653;

Best Local Similarity 75.5%; Pred. No. 3,9e-212;  
 Matches 493; Conservative 37; Mismatches 58; Indels 65; Gaps 2;

QY 1 MEARVVAALQKROVSLDCELVGSMAGAPLRYFAAEFERGTFLANLADLGVEELS 60  
 DB 1 MEARVVAALQKROVSLDCELVGSMAGAPLRYFAAEFERGTFLANLADLGVEELS 60  
 QY 61 ARGCRVSDPTGFLINPLTDDLLNEKLDREELCGPPECVLPFQILLEKPFQIFRAE 120  
 DB 61 ARGCRVSDPTGFLINPLTDDLLNEKLDREELCGPPECVLPFQILLEKPFQIFRAE 120  
 QY 121 LWRVDINDSPVFLDREITLNTLESTPGATFLLESAPDSGVINLNTSSVYVPHI 180  
 DB 121 LWRVDINDSPVFLDREITLNTLESTPGATFLLESAPDSGVINLNTSSVYVPHI 180  
 QY 181 NVHNDGEGNVSELYDKVLDREVEPELRLTLTGDDGSPRSGTLLIRLIVLDINDNP 240  
 DB 181 NVHNDGEGNVSELYDKVLDREVEPELRLTLTGDDGSPRSGTLLIRLIVLDINDNP 240  
 QY 241 EFVESLYKQVPENSPVGLVTVTSARDLDSNGEIIYAFYATERLTKFRINSTSGN 300  
 DB 241 EFVESLYKQVPENSPVGLVTVTSARDLDSNGEIIYAFYATERLTKFRINSTSGN 300  
 QY 301 LHLKAEVYEAIGYTLTIOAKDGGGLSGKCTVVVHVVDINDNPPELLMSLTSPFENS 360  
 DB 301 LHLKAEVYEAIGYTLTIOAKDGGGLSGKCTVVVHVVDINDNPPELLMSLTSPFENS 360  
 QY 361 PETVVAVERIRDRSGNNAKVCISIQHLPLVLRKSVENFTVTERALDREERTENIT 420  
 DB 361 PETVVAVERIRDRSGNNAKVCISIQHLPLVLRKSVENFTVTERALDREERTENIT 420  
 QY 421 ITVTDLGTPLRKTQNLVTVSDVNDNAPFESQTYTLRVENNSPLHIGSVATDRDS 480  
 DB 421 ITVTDLGTPLRKTQNLVTVSDVNDNAPFESQTYTLRVENNSPLHIGSVATDRDS 480  
 QY 481 GANA----- 484  
 DB 481 GTMAOVISLPSODPHPLPLASLVSINADNGHLFALSPWTMPLRAPFRNNSPALHIGS 540  
 QY 485 -----QVTYSLPLPHDPQLPGSLVSTNADNGOLFALSRDLEALQAFERRVGAADR 536  
 DB 541 VSTAETNIGVYISLPPRNPPLPLASLVSINTDNGHLFALSRDLEALQAFERRVGAADR 600

QY 537 GSPALSSQALVRLVADANDNAPFLVPLQNGSAPCTELVPR-AAAGYLVAK 588  
 DB 601 GSTALSSQALVRLVADANDSSLFVLPPLQNGSAPCTELMRAAERGGLVTK 653

RESULT 5  
 ABG17157  
 ID ABG17157 standard; Protein: 797 AA.  
 AC ABG17157:  
 DT 18-FEB-2002 (first entry)  
 DE Novel human diagnostic protein #17148.  
 KW Human: chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 OS Homo sapiens.  
 PN WO200175067-A2.  
 PD 11-OCT-2001.  
 PF 30-MAR-2001; 2001WO-US08631.  
 PR 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 PA (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;  
 XX WPI: 2001-639362/73.  
 DR N-PSDB: AAS81344.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity  
 XX  
 PS Claim 20; SEQ ID No 47516; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 797 AA:

Query Match 71.8%; Score 2304; DB 22; Length 797;  
 Best Local Similarity 71.3%; Pred. No. 8e-197;  
 Matches 447; Conservative 77; Mismatches 103; Indels 0; Gaps 0;

QY 1 MEARVHAHQKROVSLCFLGVSMAGAEPLRYFAEETERTFLANLAIDGLGVEELS 60  
 DB 1 MEIRGALDRKROVLFVLVLGLSAGTESAHYVAEETIGSFVANLARDELGLVEELS 60  
 QY 61 ARGCRIVSDEPTIGFLNPLTGLDLNLEKIDREELCGPTPEVLPFOLLLEKFOIFPAE 120  
 DB 61 SREARVSDDNKKYHLDLTLGNLLNLEKIDREELCGSTPEVLPFOLLLEKFOIFPAE 120  
 QY 121 LWRVDINDHSPVFLDREITLNLLESTTPGATFLLESANDSDGINNLNRYTSSNVYFHI 180  
 DB 121 LCVKIDINDHSPFLDKETIKISGETVAGATFLMESADLDGNSLONRYTSSNVYFHI 180  
 QY 181 NVHNDGEGVYSEIVLADKLDREVEPELRLTLGLDGGSPRSCTTLIRILVLDINDVNP 240  
 DB 181 KIPDSSDKRIYELVLDRLDLEQEEKELTLTAAVDGSPRSCTTLIRILVLDINDVNP 240  
 QY 241 EEVESLYKQVDPENSPVGLVTVTSARDLDTGSGNEIYAFYAFERTLKTPIRINSGN 300  
 DB 241 EEPQSLIEQVDPEDRLGSMWITISAKDLADAGNKGKISTEPHASEDIRKTEINPISGE 300  
 QY 301 LHLKAELEAYIQYTLTIOAKDGGGLSGCTVVVAVHTDINDNPELIMSLTSPIPENS 360  
 DB 301 VNLRSPLDEVIQSYTINIQATDGGGLSGCTLLVAVMDINDNPEVITSTIKRIPERA 360  
 QY 361 PEYVAVERIRDRDGSNNKAMVCSIODHLPEYLAKEVNEFYTLVERALDREERTYNT 420  
 DB 361 SETLVALFSLIDDSGDNRMICSIODNLPFLKPFKNFPLVSEKALDRSQAQYNT 420  
 QY 421 IYVTDGEPRLKQHNLTAVSDVNDNAPTFSGQYTYTLVRNNSPALHIGSVSATDRDS 480  
 DB 421 IYVTDGEPRLKTEYNTIYLLSDVNDNAPTFQSTITTLVRNNSPALHIGSVSATDRDS 480

QY 481 GANAQVYTSLLPPHDPOLVGLSVTSINADNGOLFALRSIDFEALQAFERRVGAADRGSPA 540  
 DB 481 GTNAQVYTSLLPPQDRHLPLASLVTSINADNGHFLALRSIDYEALQAFERRVGAADRGSPA 540  
 QY 541 LSSQALVRVLVADANANAPVLVPLONGSAPCELVPRAAEGYLVAKVAVADGSGQNA 600  
 DB 541 LSSALVRVLVADANDLPVLVPLONGSAPCELVPRAAEGYLVAKVAVADGSGQNA 600  
 QY 601 WLSYQLKATEPGLFGVWANGGEVRTA 627  
 DB 601 WLSYQLKATEPGLFGVWANGGEVRTA 627

RESULT 6  
 AAR86865  
 ID AAR86865 standard; Protein; 787 AA.

AC AAR86865;

DT 27-AUG-1996 (first entry)

DE Human protocadherin pc3.

KW Protocadherin; pc3; pc4; pc5; human; rat; cadherin; cell adhesion;

KW catenin; therapy.

OS Homo sapiens.

PN W09600289-A1.

PD 04-JAN-1996.

PE 26-JUN-1995; 95WO-0508071.

PR 27-JUN-1994; 94US-0268161.

PA (DOHE-) DOHENY EYE INST.

PI Suzuki S;

DR WPI: 1996-068873/07.

DR N-PSDB: AAT03572.

PS Claim 15; Page 115-119; 146pp; English.

CC AAR86865-R86867 represent the sequences for three protocadherins. This

CC sequence represents the human protocadherin pc3. These sequences are

CC related to cadherin, and possess cell adhesive ability. Cadherins are

CC glycosylated integral membrane proteins that are involved in cell-cell

CC adhesion. Cadherins are composed of an N-terminal extracellular domain

CC which consists of 5 unique subdomains, a membrane spanning domain, and a

CC C-terminal cytoplasmic domain. The cytoplasmic domain interacts with

CC the cytoskeleton through catenins and other cytoskeleton associated

CC proteins. The cytoplasmic domain is not present in all cadherins, but in

CC those which possess it, it is essential for the cadherins adhesive

CC function. The cadherins which do not possess a cytoplasmic domain appear

CC to function via a different method from those with a cytoplasmic domain.

CC These sequences were isolated using primers 1 and 2 (see AAT03575 and

CC AAT03576) The proteins may have regulatory functions in the cell, as

CC well as the cell-cell adhesive properties. Antibodies produced against

CC these sequences are useful for modulating the binding a

CC protocadherins, and can be used therapeutically.

XX Sequence 787 AA;

Query Match 71.3%; Score 2287; DB 17; Length  
 Best Local Similarity 72.3%; Pred. No. 2.6e-195;  
 Matches 447; Conservative 67; Mismatches 104; Indel

QY 10 OKROVSLCFLGVSMAGAEPLRYFAEETERTFLANLAIDGLGVEELS

```

Db      10 EOROVLLILLLEVLAGEPRRYSYMETERGSEFVAMLANDLIGVGLAERGAKEVSE 69
Qy      70 FTIGFLLNPLTLLNENKLDRELKGPTEPCVLPOLLLEKPKQTRALMWRDNDH 129
Db      70 DNEGQLDLDLQGLILNEKLDREKLCGTEPCIMHFVLLKPLEVFAELVTDNDH 129
Qy      130 SVEFLDREITLLESTPGATFELLESADSDVGINLRNNTYSSNVYFHHVNDNGEN 189
Db      130 SEEPFEREMTKLIPETSSIGTFVFLKARDDVGNNTYSSNVYFHHVNDNGEN 189
Qy      190 VSEVLVLDKVDREVEPELRLTLGLDGGSPRRSGTTLRLVLDINDNVEVESLYKV 249
Db      190 KPELVLDTELDREQAELRLTLAVDGGSPRRSGTVOILLVLDANNAPEFVALYEV 249
Qy      250 QVPENSPVGLVTVASADLDGTSNGELVAFVAFVATERLTLPFRINSGMLHKAELNY 309
Db      250 QVPENSPVGLVTVASADLDGTSNGELVAFVAFVATERLTLPFRINSGMLHKAELNY 309
Qy      310 EAIQTYTLTLQAKDGGGLSKCTVYVHTVDINDNPELLMSLTSPENSPETVAVER 369
Db      310 ETMSYDIDIRASDGGGLSKCTVYVHTVDINDNPELLMSLTSPENSPETVAVER 369
Qy      370 IRDRSGNNAKWCOSIOHLPVLPKPSYENFTVTERALDRERTYNTTITVDLCTP 429
Db      370 IRDRSGNNAKWCOSIOHLPVLPKPSYENFTVTERALDRERTYNTTITVDLCTP 429
Qy      430 RLKTOHNLTVTVSDVNDNAPTESOTYTLIRVRENNSPALHIGSADRDGSGANAQVYS 489
Db      430 RLKTOHNLTVTVSDVNDNAPTESOTYTLIRVRENNSPALHIGSADRDGSGANAQVYS 489
Qy      490 LIPRDPOLPLGLSVINADNGQLFALRSLDFEALQAFERFVGAADRGSPALSSQALVR 549
Db      490 LIPRDPOLPLGLSVINADNGQLFALRSLDFEALQAFERFVGAADRGSPALSSQALVR 549
Qy      550 LVANDNAPFVLYPLONGSAPCTELVPRAEAGYLVAKVAVVGDGSGONAMLSYOLTKA 609
Db      550 LVANDNAPFVLYPLONGSAPCTELVPRAEAGYLVAKVAVVGDGSGONAMLSYOLTKA 609
Qy      610 TEPGLFGVMAHNGEVRTA 627
Db      610 TEPGLFGVMAHNGEVRTA 627

RESULT 7
AAB18920 standard; Protein; 800 AA.
AAB18920;
AAB18920;
08-FEB-2001 (first entry)

A novel polypeptide designated PRO4352.

Secreted protein; transmembrane protein; PRO1484; PRO4334; PRO1122;
PRO1889; PRO1889; PRO1885; PRO4353; PRO4357; PRO4405; PRO4556;
PRO4352; PRO4380; PRO4354; PRO4408; PRO5737; PRO4425; PRO5990; PRO6030;
PRO4424; PRO4422; PRO4430; PRO4459; tumour; obesity; diabetes;
Insulinemia; kidney disorder; Bergers disease; nephropathy;
Schonlein-Henoch purpura; celliac disease; dermatitis herpetiformis;
Crohns disease.
KW
KW
XX
XX
OS Homo sapiens.
XX
XX
Key Location/Qualifiers
XX Peptide 1..26
FT /note= "signal peptide"
FT Modified-site 22..28
FT /note= "N-myristoylation site"
FT Modified-site 27..33
FT /note= "N-myristoylation site"
FT Modified-site 53..59
FT /note= "N-myristoylation site"

```

```

FT Modified-site 82..88
FT /note= "N-myristoylation site"
FT Modified-site 162..168
FT /note= "N-myristoylation site"
FT Modified-site 169..173
FT /note= "N-glycosylation site"
FT Modified-site 181..185
FT /note= "N-glycosylation site"
FT Modified-site 184..190
FT /note= "N-myristoylation site"
FT Modified-site 217..223
FT /note= "N-myristoylation site"
FT Modified-site 324..330
FT /note= "N-myristoylation site"
FT Modified-site 325..331
FT /note= "N-myristoylation site"
FT Modified-site 394..402
FT /note= "tyrosine kinase phosphorylation site"
FT Modified-site 418..422
FT /note= "N-glycosylation site"
FT Modified-site 436..440
FT /note= "N-glycosylation site"
FT Modified-site 471..477
FT /note= "N-myristoylation site"
FT Modified-site 567..571
FT /note= "N-glycosylation site"
FT Modified-site 568..574
FT /note= "N-myristoylation site"
FT Modified-site 578..585
FT /note= "tyrosine kinase phosphorylation site"
FT Domain
FT 687..711
FT /note= "transmembrane domain"
FT Modified-site 759..765
FT /note= "N-myristoylation site"
FT Modified-site 781..785
FT /note= "amidation site"
FT Modified-site 788..792
FT /note= "N-glycosylation site"
FT WO200056889-A2.
FT 28-SEP-2000.
FT 01-MAR-2000; 2000WO-US05601.
FT 23-MAR-1999; 99US-0125774.
FT 23-MAR-1999; 99US-0125778.
FT 24-MAR-1999; 99US-0125826.
FT 31-MAR-1999; 99US-0127035.
FT 05-APR-1999; 99US-0127035.
FT 21-APR-1999; 99US-0130359.
FT 27-APR-1999; 99US-0131270.
FT 27-APR-1999; 99US-0131272.
FT 27-APR-1999; 99US-0131291.
FT 04-MAY-1999; 99US-0132371.
FT 04-MAY-1999; 99US-0132379.
FT 04-MAY-1999; 99US-0132383.
FT 25-MAY-1999; 99US-0135750.
FT 08-JUN-1999; 99US-0138166.
FT 20-JUN-1999; 99US-0144791.
FT 03-AUG-1999; 99US-0146970.
FT 09-DEC-1999; 99US-0170262.
FT (GETH ) GENENTECH INC.
FT Desnoyers L, Eaton DL, Goddard A, Godowski PJ, Gurney AL, Pan J;
FT Stewart TA, Watanabe CK, Wood WI, Zhang Z;
FT WPI; 2000-628263/60.
FT N-PSDB; AAA96347.
FT Novel secreted and transmembrane polypeptides useful for diagnosing
FT tumour in a mammal, for identifying agonists and antagonists of the

```



PT polypeptide and for therapeutic use  
XX  
PS Claim 12; Fig 24; 222pp; English.

XX The present sequence represents a secreted or transmembrane polypeptide.  
CC The specification describes polypeptides designated PRO1484, PRO4334,  
CC PRO1122, PRO1889, PRO1890, PRO1887, PRO1853, PRO4405,  
CC PRO4356, PRO4380, PRO4354, PRO4408, PRO5737, PRO4455, PRO5990,  
CC PRO6030, PRO4424, PRO4422, PRO4430 and PRO4499. PRO1889 polypeptide is  
CC useful for diagnosing tumour in a mammal. The polypeptides, their  
CC agonists and antagonists are useful treating a condition associated with  
CC expression or activity of the polypeptide. Conditions treated include  
CC obesity, diabetes or hyper- or hypo-insulinemia. The polypeptides are  
CC capable of inducing proliferation of mammalian kidney mesangial cells  
CC and are therefore useful for treating kidney disorders associated with  
CC decreased mesangial cell function such as Berger's disease or other  
CC nephropathies associated with Schönlein-Henoch purpura, celiac disease,  
CC dermatitis herpetiformis or Crohn's disease. The nucleic acids may be used  
CC to generate transgenic animals for use in development and screening of  
CC therapeutically useful reagents and also for chromosome identification  
CC and tissue typing.

XX Sequence 800 AA:

Query Match 70.1%; Score 2249; DB 21; Length 800;  
Best Local Similarity 70.4%; Pred. No. 6,7e-192;  
Matches 435; Conservative 70; Mismatches 113; Indels 0; Gaps 0;

QY 10 QKQVSLLCVFLGVSNAGAEPLRYFAETERTGFLANALDGLGVELLSRGCKRYSD 69  
DB 10 RQKQVLEFLFMGVSNLAGSGFGRYSVETETKSGFVYNAKDLGAKSLAARGTRVSD 69  
QY 70 ETIGFLNPLDGLDILNKLDEKLCPTPECVLPOLLKPPQIAELWVKLINH 129  
DB 70 DNKQYLLLSHGLNLTNKLDEKLCGRKPEMLYFQILMDPPQIYAEIRVNDINH 129  
QY 130 SPVFLDREITLNTLESTPGATFLLESANDSVGINNLNNTYISSNVYFHNWHDGECN 189  
DB 130 APFQDKKEYLYKISENTAGTAFLERADDPGLGNGIGNYITSPSEFHINISGDECM 189  
QY 190 VSELYLVKVDREYVELRLITGDDGSPPSGTTLRILYLDINENVPFESLYKYV 249  
DB 190 IYELVLDVADREDEGELESLTLTALDGSFSPKSGSTVRIYVLDVNDAPAFAGLYET 249  
QY 250 QVENSVPVGLVYVTSARLDITGSGEIVYAFYATERTLKPRINSTSGNLKRELNY 309  
DB 250 QAPENSPILGLIYKVAEDVDVSVNAEVSFSDASENIRITFOINPFGELFLRELIX 309  
QY 310 EAIQYTLTIOAKDGGGLSGCTVVVYVTDINDNPELLMSLTSPDENSPETVAVFR 369  
DB 310 ELVNSYKINIQAMDGGGLSARCVLEVLDITNDNPELLVSSFSNVAENSPETPLAVFR 369  
QY 370 IRRDGSNNNAKWCSTODHLPFLKPSVENFTYLVTERALDREKREYVNTITVPLGTP 429  
DB 370 INRDSENGKWCYIOENLPFLKPSVENFTYLLITEGALDREIRAEYNTITVTDLGP 429  
QY 430 RLKTOHNLVYVSDVNDNAPFTSCOTYTLRVENNSPALHIGSVSUTDRSGANAQVTS 489  
DB 430 RLKTEHNTIYLVSDVNDNAPAFQTSYTLFVRNNSPALHIGSVSUTDRSGNNAQVTS 489  
QY 490 LLEPHDPLDGLSVLSINADNGQLFALSLDFEALQAFERRVAGADRGSFALSSQALVRV 549  
DB 490 LLEPHDPLDGLSVLSINADNGHLFALSLDYBALQAFERRVAGADRGSFALSSQALVRV 549  
QY 550 LVADANDNAFVLYPLONGSAPCTELVPRAAEGYLVAKVYVADGSGQANMISTYLLKA 609  
DB 550 LVADANDNSPFLVPLONGSAPCTELVPRAAEGYLVAKVYVADGSGQANMISTYLLKA 609  
QY 610 TERGLFGVMAHNGEVRTA 627  
DB 610 TERGLFGVMAHNGEVRTA 627

RESULT 8  
ID AAU29262  
AAU29262 standard; Protein; 800 AA.  
AC AAU29262;  
XX 18-DEC-2001 (first entry)  
DE Human PRO polypeptide sequence #239.  
XX  
KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;  
KW dog; cat; pig; goat; rabbit; tumour; necrosis factor alpha; TNF-alpha;  
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;  
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN W0200168848-A2.  
XX  
PD 20-SEP-2001.  
XX  
PE 28-FEB-2001; 2001WO-US06520.  
XX  
PR 01-MAR-2000; 2000WO-US05601.  
PR 02-MAR-2000; 2000WO-US05841.  
PR 03-MAR-2000; 2000US-187202P.  
PR 06-MAR-2000; 2000US-186968P.  
PR 14-MAR-2000; 2000US-189320P.  
PR 14-MAR-2000; 2000US-189328P.  
PR 15-MAR-2000; 2000WO-US06884.  
PR 21-MAR-2000; 2000US-190828P.  
PR 21-MAR-2000; 2000US-191007P.  
PR 21-MAR-2000; 2000US-191048P.  
PR 21-MAR-2000; 2000US-191314P.  
PR 28-MAR-2000; 2000US-192655P.  
PR 29-MAR-2000; 2000US-193032P.  
PR 29-MAR-2000; 2000US-193033P.  
PR 30-MAR-2000; 2000WO-US08439.  
PR 04-APR-2000; 2000US-194449P.  
PR 04-APR-2000; 2000US-194647P.  
PR 11-APR-2000; 2000US-195975P.  
PR 11-APR-2000; 2000US-196000P.  
PR 11-APR-2000; 2000US-196187P.  
PR 11-APR-2000; 2000US-19690P.  
PR 11-APR-2000; 2000US-196820P.  
PR 18-APR-2000; 2000US-198121P.  
PR 18-APR-2000; 2000US-198585P.  
PR 18-APR-2000; 2000US-198397P.  
PR 25-APR-2000; 2000US-199550P.  
PR 25-APR-2000; 2000US-199654P.  
PR 03-MAY-2000; 2000US-201516P.  
PR 17-MAY-2000; 2000WO-US13705.  
PR 22-MAY-2000; 2000WO-US14042.  
PR 30-MAY-2000; 2000WO-US14941.  
PR 02-JUN-2000; 2000WO-US15264.  
PR 05-JUN-2000; 2000US-209832P.  
PR 28-JUL-2000; 2000WO-US20710.  
PR 22-AUG-2000; 2000US-0644848.  
PR 24-AUG-2000; 2000WO-US23328.  
PR 08-NOV-2000; 2000WO-US30952.  
PR 01-DEC-2000; 2000WO-US32678.  
PR 20-DEC-2000; 2000WO-US34956.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;  
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;  
XX  
XX WPI; 2001-602746/68.  
XX N-PSDB; AAS46163.  
PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the  
presence of tumours, such as prostate and breast tumours, in mammals and

to screen for modulators of the compounds -

Claim 11; Fig 478; 774pp; English.

Sequences AAU29024-AAU99378 represent PRO polypeptides of the invention. The PRO polypeptides and their associated nucleic acids can be used to detect the presence of a tumour in a mammal by comparing the level of expression of a PRO polypeptide in a test sample of cells from the animal and a control sample of normal cells, whereby a higher level of expression in the test sample indicates the presence of a tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats and rabbits but are preferably human. The polypeptides can be used to stimulate tumour necrosis factor (TNF) alpha release from human blood, when contacted with it. A specific polypeptide can be used to stimulate the proliferation or differentiation of chondrocyte cells. The PRO proteins can be used to determine the presence of tumours and also susceptibility to tumour development, particularly adrenal, lung, colon, breast, prostate, rectal, cervical, or liver tumours, in mammalian subjects. The oligonucleotide probes specific for the PRO nucleic acids can be used for genetic analysis of individuals with genetic disorders.

Sequence 800 AA;

Query Match 70.1%; Score 2249; DB 22; Length 800;  
Best Local Similarity 70.4%; Pred. No. 6.7e-192;  
Matches 435; Conservative 70; Mismatches 113; Indels 0; Gaps 0;

QY 10 OKROVSLCYFLGVSMGAEPLRYFAEETRGFTFLANLADLGLVEELSGRCRIVSD 69  
DB 10 ROROVLEFLFVGSLAGSGFRYSVTEERKGSFVNNAKDLGANGELAAAGTRVSD 69  
QY 70 ETIFLLNPLTGDLINLNEKIDRELCGPTPCVLPOLLLEKPFQFRAELVWROINDH 129  
DB 70 DNKOYLIDSHGTGWLNEKIDRELCGPKPEPCMLYFOIIMDPFOYLAELVWROINDH 129  
QY 130 SPVLDREITLNESTTPGATFLLESADSDYGINNLRVYTTISNVEFINNHDCEN 189  
DB 130 APVQDEYVAKISENABEGTARERADPDGGLNQYTTISPSFHTINSGDEGN 189  
QY 190 VYSLVLDKVLDEPELRLTLTGDSPPRSGLTLRLIWLIDINVPFVESLYKY 249  
DB 190 IYELVLDKALDREOGELSLTLTALDGSPPRSGLSTVRIVLDVNDAPQAQALYET 249  
QY 250 QVPSNVSGLVYVSARDLDTGSGNBIYAFYATERTLKTFRINSGNLHKLKELNY 309  
DB 250 QAPNSPGLIYKVAWADVDGVAENVSSEFDASENIRTTQIINPFGSEIFRELIDY 309  
QY 310 EAIQTYTLTQARDGGLSGCTVYVVDINDNPPELLMSLTPSPENSPETVAVER 369  
DB 310 ELVNSYKINIOAMDGGGLSARCRIVEYLDTNPELIVSSNSVANSPEPTPLAVK 369  
QY 370 IRRDGSNNKAKMCSIODHLPEVLKTSVENFYTLVTERAIDREERTVNTTVDLGP 429  
DB 370 INRDSENGKWCYIOENLPFLKPSVENFYTLITEGALDREIRAEVNTTITVDLGP 429  
QY 430 RLKTOHNLVTVSDVNDNAPTFQSTYTLVRENNSPALHIGSVSTDDSGANQVYS 489  
DB 430 RLKTEHNTITVSDVNDNAPTFQSTYTLVRENNSPALHIGSVSTDDSGANQVYS 489  
QY 490 LLEPHDPLPLGSLVSNADNGQLFALRSIDFEALQAFERRVGAADRGSPALSOALYV 549  
DB 490 LLEPHDPLPLASLVSINADNGHLFALRSIDYALQAFERRVGAATRGSPALSRALYV 549  
QY 550 LVADADNAPFVLYPLQNSAPCTELVPRAAEGYLVAAVAVDGDGSGANMLSYOLKA 609  
DB 550 LVADADNAPFVLYPLQNSAPCTELVPRAAEGYLVAVVAVDGDGSGANMLSYOLKA 609  
QY 610 TEPGLGVAAHNGEVRTA 627  
DB 610 TEPGLGVAAHNGEVRTA 627

RESULT 9

ABG23214  
ID ABG23214 standard; Protein; 1888 AA.

AC ABG23214;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #23205.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PDB: AAS87401.

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity

Claim 20; SEQ ID No 53573; 103pp; English.

The invention relates to isolated polynucleotide (I) and  
polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC and responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pcl\_sequences.

Sequence 1888 AA;

Query Match 69.7%; Score 2237; DB 22; Length 1888;  
Best Local Similarity 71.1%; Pred. No. 3.3e-190;  
Matches 441; Conservative 69; Mismatches 108; Indels 2; Gaps 2;

QY 10 OKROVSLCYFLGVSMGAEPLRYFAEETRGFTFLANLADLGLVEELSGRCRIVSD 69  
DB 26 EOROVLLTLTGLTGLWGEPRYSVMEERKGSFVNNAKDLGANGELAAAGTRVSD 85  
QY 70 ETIFLLNPLTGDLINLNEKIDRELCGPTPCVLPOLLLEKPFQFRAELVWROINDH 129  
DB 86 DNEGGLDLDGTGLIINLNEKIDRELCGPTPCVLPOLLLEKPFQFRAELVWROINDH 145



BR 11-SEP-1998; 98US-0100038.  
 XX (GETH ) GENENTECH INC.  
 XX  
 XX  
 PI Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;  
 XX  
 XX WPI: 1999-551358/46.  
 DR N-PSDB: AA234233.  
 XX  
 XX New secreted and transmembrane polypeptides and their polynucleotides,  
 PT useful for treating blood coagulation disorders, cancers and cellular  
 PT adhesion disorders -  
 XX  
 PS Claim 12; Fig 163; 530pp; English.  
 XX  
 XX The present invention describes secreted and transmembrane polypeptides  
 CC and their polynucleotides. The nucleotide sequences are useful as  
 CC sources of probes, primers, for chromosome mapping, and for generation  
 CC of antisense sequences. They can also be used to create transgenic  
 CC animals. The proteins can be used to treat a variety of diseases and  
 CC disorders, depending on their function. Diseases that may be treated  
 CC include blood coagulation disorders, cancers and cellular adhesion  
 CC disorders. They may also be used to raise antibodies. AA233891 to  
 CC AA23438, and AA41685 to AA41774 represent polynucleotide and  
 CC polypeptide sequence given in the exemplification of the present  
 CC invention.  
 XX  
 XX Sequence 798 AA;  
 SQ  
 Query Match 66.8%; Score 2142; DB 20; Length 798;  
 Best Local Similarity 68.0%; Pred. No 2.5e-182;  
 Matches 427; Conservative 66; Mismatches 133; Indels 2; Gaps 2;  
 1 MEANVAHALOKROVSLLCVFLGSMAG-ABPLRYFAEETGERGFLAMLAIDGLGVEEL 59  
 1 MEASGKLICRQVLFSLFLLGLSLAGAEPRYSVEETGESFYNLAKDGLGREF 60  
 60 SARCRIVSDETIGFLLNPLTGLLNEKIDRELCGPFEPVLPQLLEKPFOTFRA 119  
 61 SRGRVAVSRKGNKHLQLOMETADLLNEKIDRELCGHPTECVLRQVLESPPFEFOA 120  
 120 ELWRDNDHSPVLDREITLNLLESTTPGATFELLESABSDGGINLRYVTSSNYPH 179  
 121 ELQYIDNDHSPVLDKQMLVYKVSSESPPTTPKMAELDVGQNNIENYITSPSYFR 180  
 180 INVHNGEGNVSELYLDKVLDRREVEPLRLTLGLDGGSPRSGTTLRIIVLDINDY 239  
 181 VTRKRSDGKRYPPELVLDKALDREVEALRLTLGLDGGSPRSGTAQVYIEVLVDNDA 240  
 240 PEPFESLYKQVSPNSPVSLVNTVSARDLDGNSGFIYAFYATFERTIKTRINSTG 299  
 241 PEFQPTFRQVISEDSPVGLVAVKVSATDVDTGNGEISYSLFQASEEIGTKFKINPLTG 300  
 300 NLHLALNEALQIOTYTLTQAKDGGISGCTVYVHTDINDNPELMSLTSPFIPEN 359  
 301 EIELKQLODEKIQSYEVNIEARDAGTFSKCTVLIQVIDVNDHADEVMSATTSIPFEN 360  
 360 SPETVAVAFIRDRDSGNNAKWCISODHLPVLAUKSVENFYTLVTERALDREERTYNI 419  
 361 APETVVALFESVLDGSENGKISCSIOEDLPFLK-SAEKNFYTLTERPLDRSRAEYNI 419  
 420 TIIVTDGTPRLKTOHNLITVTSVDVNDNAPTFSQTYTLRVRNNNSPALHIGSVATDND 479  
 420 TIIVTDGTPRLKTOHNLITVTSVDVNDNAPTFSQTYTLRVRNNNSPALHIGSVATDND 479  
 480 SGANAQVTSILPDPHPOLPLGLSVGINADNGQLFALRSIDFALDAFEFRVGAARSP 539  
 480 SGANAQVTSILPDPHPOLPLGLSVGINADNGQLFALRSIDFALDAFEFRVGAARSP 539  
 540 ALSQALVRVLVADANDAPFVLYPLQNSAPCTELVPRAAEGYLVARVAVDGGSON 599  
 540 ALSQALVRVLVADANDAPFVLYPLQNSAPCTELVPRAAEGYLVARVAVDGGSON 599

QY 600 AMLSYQLKATEPGLFGVAHNGEYRTA 627  
 DB 600 AMLSYQLKATEPGLFGVAHNGEYRTA 627  
 RESULT 11  
 AAB44302  
 ID AAB44302 standard; Protein; 798 AA.  
 XX  
 XX AAB44302;  
 AC  
 XX 08-FEB-2001 (first entry)  
 DT  
 XX Human PRO531 (UNQ332) protein sequence SEQ ID NO:405.  
 DE  
 XX Human; secreted protein; transmembrane protein; PRO; EST; cytosolic;  
 KW expressed sequence tag; detection; cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200053756-A2.  
 XX  
 PD 14-SEP-2000.  
 PF 18-FEB-2000; 2000WO-US04341.  
 XX  
 XX 08-MAR-1999; 99WO-US05028.  
 PR 12-MAR-1999; 99US-0123957.  
 PR 29-MAR-1999; 99US-0126773.  
 PR 21-APR-1998; 99US-0130232.  
 PR 28-APR-1999; 99US-0131445.  
 PR 14-MAY-1999; 99US-0134287.  
 PR 23-JUN-1999; 99US-0141037.  
 PR 26-JUL-1999; 99US-0145698.  
 PR 29-OCT-1999; 99US-0162506.  
 PR 30-NOV-1999; 99US-0528313.  
 PR 02-DEC-1999; 99WO-US28551.  
 PR 02-DEC-1999; 99WO-US28565.  
 PR 16-DEC-1999; 99WO-US30095.  
 PR 30-DEC-1999; 99WO-US31243.  
 PR 30-DEC-1999; 99WO-US31274.  
 PR 05-JAN-2000; 2000WO-US00219.  
 PR 06-JAN-2000; 2000WO-US00277.  
 PR 06-JAN-2000; 2000WO-US00376.  
 XX  
 XX (GETH ) GENENTECH INC.  
 XX  
 XX Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL,  
 PI Ferraira N, Filvaroff E, Fong S, Gerber H, Gerritsen ME,  
 PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ,  
 PI Kljavin IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA,  
 PI Shelton DL, Stewart TA, Tumas D, Williams PM, Wood WI;  
 DR WPI: 2000-611443/58.  
 DR N-PSDB: AAC78558.  
 XX  
 XX Novel PRO polypeptides and polynucleotides used in detection methods,  
 PT to target bioactive molecules to specific cells, and to modulate  
 PT cellular activities -  
 XX  
 PS Claim 12; Fig 163; 636pp; English.  
 XX  
 XX AAC78458 to AAC78599 represent polynucleotide and EST (expressed  
 CC sequence tag) sequences which encode secreted or transmembrane PRO  
 CC polypeptides. The PRO polynucleotides and polypeptides have cytostatic  
 CC activity. The polynucleotides and polypeptides can be used for detecting  
 CC the presence of PRO polypeptides in samples, for linking bioactive  
 CC molecules to cells and for modulating biological activities of cells,  
 CC using the polypeptides for specific targeting. The polypeptide targeting  
 CC can be used to kill the target cells, e.g. for the treatment of cancers.  
 CC The polypeptide pairs provide specific targeting of bioactive molecules  
 CC to cells. AAC78600 to AAC78987 represent PCR primers and probes used in  
 CC the isolation of the PRO polynucleotide sequences.



```

Db 301 EIELKQDLEFKLOSIEYENIEARDAGTFSGKCYVLIVDNDHAPVNTSAFTSPLEN 360
QY 360 SPETVAVAFRIIRDSDGNNAKMCSDIDHLPFLVKPSEVNEFYTLVTRALDREERTFYNI 419
Db 361 APEITVALFVSDDSDSGENKISCSIEDLPEFLK -SAENFYTLITLRIPIDRESRAYNI 419
QY 420 TITVTDGTPRLKQHNILTVSVVDNNAFPFSSOTYTTLVRENNSFALHIGVSATDRD 479
Db 420 TITVTDGTPMLITQLMKMTIADVNNAPAFOTSTYTLFVRENNSFALHIRSVSATDRD 479
QY 480 SGANAQYTSLSLPPHDPOLPGSLVSNADNGOLFALRSLDPEFLAQEFERVGAAADRGSP 539
Db 480 SGTMAQYTSLSLPPHDPOLPGSLVSNADNGHFLALRSLDPEFLAQEFERVGASDHGSP 539
QY 540 ALSSQALVRLVADANDNAFVLYPLQNGSAPCTELVPAAGVYLVAKVAADGSGON 599
Db 540 ALSSQALVRLVADANDNAFVLYPLQNGSAPCTELVPAAGVYLVAKVAADGSGON 599
QY 600 AMLSYQLLKATPEGLFEGVMAHNEVFTA 627
Db 600 AMLSYQLLKATPEGLFEGVMAHNEVFTA 627

```

## RESULT 13

AAB33433 standard; Protein; 798 AA.

AAB33433:

29-JAN-2001 (first entry)

Human PRO531 protein UNQ332 SEQ ID NO:107.

Human; immune related disease; diagnosis; antinflammatory; cardiant;  
 dermatologic; antiarthritis; antirheumatic; immunosuppressive;  
 haemostatic; antithyroid; antidiabetic; neurotropic; neuroprotective;  
 antiaesthetic; hepatotropic; virucide; antiparasitic; antiallergic;  
 osteoarthritis; spondyloarthritis; systemic sclerosis; sarcoidosis;  
 idiopathic inflammatory myopathy; Sjogren's syndrome; thyroiditis;  
 systemic vasculitis; autoimmune haemolytic anaemia; diabetes mellitus;  
 autoimmune thrombocytopenia; immune-mediated renal disease;  
 demyelinating disease; hepatobiliary disease; Whipple's disease;  
 inflammatory bowel disease; gluten-sensitive enteropathy;  
 autoimmune disease; immune-mediated skin disease; allergic disease;  
 immunological disease; transplantation associated disease;  
 graft rejection; graft-versus-host-disease.

Homo sapiens.

WO200053758-A2.

14-SEP-2000.

02-MAR-2000; 2000WO-US05841.

```

PR 08-MAR-1999; 99WO-US05028.
PR 10-MAR-1999; 99US-0123618.
PR 12-MAR-1999; 99US-0123957.
PR 23-MAR-1999; 99US-0123775.
PR 12-APR-1999; 99US-0128849.
PR 20-APR-1999; 99WO-US08615.
PR 28-APR-1999; 99US-0131445.
PR 04-MAY-1999; 99US-0133371.
PR 14-MAY-1999; 99US-0134287.
PR 02-JUN-1999; 99WO-US12252.
PR 23-JUN-1999; 99US-0141037.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 28-JUL-1999; 99US-0146222.
PR 01-SEP-1999; 99WO-US20111.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.

```

```

PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-OCT-1999; 99US-0162506.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28409.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 16-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30999.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.

```

(GETH ) GENENTECH INC.

Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W,  
 Kabakoff RC, Lu Y, Pan J, Pennica D, Shelton DL, Smith V,  
 Stewart TA, Tamas D, Watanabe CK, Wood WI, Yan M.

WPI: 2000-572271/53.

N-PSDB: AAC58598.

Sixty four PRO polypeptides, useful in the diagnosis and treatment of  
 immune related disorders, e.g. systemic lupus erythematosus, rheumatoid  
 arthritis, osteoarthritis, thyroiditis and diabetes mellitus.

Claim 33; Fig 40; 309pp; English.

The present invention describes sixty four human PRO proteins which can  
 be used in the treatment of immune related diseases. The human PRO  
 proteins, anti-PRO antibodies, agonists and antagonists are useful for  
 treating and diagnosing immune related disorders. The disorders are  
 selected from systemic lupus erythematosus, rheumatoid arthritis,  
 osteoarthritis, juvenile chronic arthritis, spondyloarthritis,  
 systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's  
 syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic  
 anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus,  
 immune-mediated renal disease, demyelinating diseases of the central  
 and peripheral nervous systems, hepatobiliary diseases, inflammatory  
 bowel disease, gluten-sensitive enteropathy and Whipple's disease,  
 autoimmune or immune-mediated skin diseases, allergic diseases,  
 immunological diseases of the lung, and transplantation associated  
 diseases including graft rejection and graft-versus-host-disease.  
 AAC58397 to AAC58578 represent PCR primers and hybridisation probes used  
 in the isolation of human PRO sequences. AAC58579 to AAC58642 and  
 AAB33414 to AAB33477 represent human PRO polynucleotide and protein  
 sequences given in the exemplification of the present invention.

Sequence 798 AA;

Query Match 66.8%; Score 2142; DB 21; Length 798;  
 Best Local Similarity 68.0%; Pred. No. 2,5e-182;  
 Matches 427; Conservative 66; Mismatches 133; Indels 2; Gaps 2;

```

QY 1 MEARVVAHAKROVSLICFLGVSMAG-APPLRYFAEETERTPTLANLAIDGLVEEL 59
Db 1 MEASGKILCRQRYLFSLLLGLSLAGAAPRSTSYVEETEGSSFTYNLAKDGLQREF 60
QY 60 SARCRIVSDETIGFLLINPLTGDLLINLEKLDREELCGPTPCVLPFOLLKEPQIFRA 119
Db 61 SRKGVRYVSRGNKLHLQINQETADLLINLEKLDREDCGHTPECVLRFQVLESPPFFQ 120
QY 120 ELWVRDINDHSVFLDREITLNLLESTTPGATFLLESANDSDVGINNLNNTYSSNVYH 179

```

```

Db      121 ELQVIDINDHSPFLDKQMLVKVSESSPGTFFPLKNAEDDVGQNNIENYIISPNYSFR 180
Qy      180 INVHNGSGNANYSELYLDKVDREVPRLRLTGLDGGSPRRSTTLIRLIVLDINDV 239
Db      181 VLTFRRSQGRKPELVLDKALDREBEALRLTLALDGGSPRRSTAAVYIEVLDVNDNA 240
Qy      240 PEFVSLKVOVPENSPGSLVTVVSARDLDTGSGEIVAFVFAVTEERTLTFRINSYG 299
Db      241 PEFEPFVRVQISEDSPPGFLVVKVSAVDVDTGVNGEISLSPFOASEIGTFKINPLTG 300
Qy      300 NLHLKAEINYEALIQTYTLTQAKDGGISGKCTVVVAVHTDINDNPELLMSLTSPIDEN 359
Db      301 EIELKKQLDFEKLQSYEVNIEARDAFTSGKCTVLIQYIDVNDHAPVETMSAFTSPIDEN 360
Qy      360 SPETVVAVFRIRDRDSSGNNAKAVCSIODHLPEVLKPSVENYTLTVERALDREKREYNI 419
Db      361 ABEIVVALEFSVSDLDGSENGKISCSIOEDLPELAK-SAENETTLTTERPLDRESAEYNI 419
Qy      420 TTTVVDLGTPLRKTQHNLTVTVDVNDNAPFESQTTTLTRVRENSPALHIGVSATDRD 479
Db      420 TTTVVDLGTPLRKTQHNLTVTVDVNDNAPFESQTTTLTRVRENSPALHIGVSATDRD 479
Qy      480 SGANNAVTVSLIPHPDPLIGSLVSIINADNGQLFALNSLDEALQAFEFVGAADRGSP 539
Db      480 SGTNAVTVSLIPHPDPLIGSLVSIINADNGQLFALNSLDEALQAFEFVGAADRGSP 539
Qy      540 ALSSALVAVLADANDNAPFLVYPLONGSAPCTELVPRAAAGVIAVKVAVADGSDGON 599
Db      540 ALSSALVAVLADANDNAPFLVYPLONGSAPCTELVPRAAAGVIAVKVAVADGSDGON 599
Qy      600 AMLSYQLKATEPGLGEGVAHNGEVRTA 627
Db      600 AMLSYQLKATEPGLGEGVAHNGEVRTA 627

```

## RESULT 14

AAU83657

ID AAU83657 standard; Protein: 798 AA.

XX AAU83657;

XX 08-MAY-2002 (first entry)

DE Human PRO protein, Seq ID No 132.

KW Human; secreted protein; PRO; tumour; lung cancer; colon cancer;

KW breast cancer; prostate tumour; rectal tumour; liver tumour;

KW pericyte cell proliferation; chondrocyte cell proliferation;

KW tumour necrosis factor-alpha.

OS Homo sapiens.

PN WO200208286-A2.

PD 31-JAN-2002.

PE 29-JUN-2001; 2001WO-US21066.

XX 20-JUL-2000; 2000US-219556P.

XX 25-JUL-2000; 2000US-220585P.

XX 25-JUL-2000; 2000US-220605P.

XX 25-JUL-2000; 2000US-220607P.

XX 25-JUL-2000; 2000US-220624P.

XX 25-JUL-2000; 2000US-220638P.

XX 25-JUL-2000; 2000US-220664P.

XX 25-JUL-2000; 2000US-220666P.

XX 26-JUL-2000; 2000US-220893P.

XX 28-JUL-2000; 2000WO-US20710.

XX 23-AUG-2000; 2000WO-US23522.

XX 24-AUG-2000; 2000WO-US23328.

XX 15-SEP-2000; 2000US-000000P.

XX 10-NOV-2000; 2000WO-US30873.

```

PR      28-NOV-2000; 2000US-253646P.
PR      01-DEC-2000; 2000WO-US32678.
PR      20-DEC-2000; 2000US-0747259.
PR      20-DEC-2000; 2000WO-US34956.
PR      28-FEB-2001; 2001WO-US06520.
PR      10-MAY-2001; 2001US-0854280.
PR      25-MAY-2001; 2001WO-US17092.
XX
PA      (GETH ) GENENTECH INC.
XX
XX      Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI      Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
DR      WPI; 2002-172001/22.
XX
XX      N-PSDB; ABR33601.
XX
XX      One hundred and twenty two nucleic acids encoding PRO polypeptides,
PT      useful for treating a PRO related disorder and for diagnosing tumours
PT      such as lung cancer, colon cancer, breast tumour, prostate tumour, rectal
PT      tumour or liver tumour.
XX
XX      Claim 11; Figure 132; 359pp; English.
XX
XX      The invention relates to one hundred and twenty two nucleic acids
XX      encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides
XX      encode human secreted proteins. The PRO nucleic acids, polypeptides,
XX      agonists and antagonists are useful for treating a PRO related disorder.
XX      The PRO polypeptides are useful for diagnosing tumours, especially lung
XX      cancer, colon cancer, breast tumour, prostate tumour, rectal tumour or
XX      liver tumour. The PRO polypeptides are useful for stimulating the
XX      proliferation of, or gene expression, in pericyte cells, for stimulating
XX      the proliferation or differentiation of chondrocyte cells, for
XX      stimulating the release of tumour necrosis factor-alpha from human dermal
XX      for stimulating or inhibiting the proliferation of normal human dermal
XX      fibroblast cells. The PRO polypeptide may also be used as molecular
XX      weight markers and for tissue typing. The PRO nucleic acids have
XX      applications in molecular biology, including use as hybridisation probes,
XX      and in chromosome and gene mapping. AAU83592-AAU83713 represent human PRO
XX      protein sequences of the invention.
XX
SO      Sequence 798 AA;
XX
Query Match 66.88; Score 2142; DB 23; Length 798;
Best Local Similarity 68.0%; Pred. No. 2,5e-182;
Matches 427; Conservative 66; Mismatches 133; Indels 2; Gaps 2;
Qy      1 MEARVVAHLQKROYSLCFLVGVSMAG-APRLRFVAFEEFERGFLNMLNIDIGLVEEL 59
Db      1 MEASGKLICRQOVLFSLFLGLSLAGAAEPRTSVVEETEGSFTVNLAKDGLDQREF 60
Qy      60 SARGCRAVSDETIGFLINPLTGDLLNEXKLDRELCPTEPCVLPQOLLLEKPFQIFRA 119
Db      61 SRGCVAVSRGNKLAHLQUNOETADLLNEXKLDRELCPTEPCVLPQOLLLEKPFQIFRA 120
Qy      120 ELWVRDINDHSPVFLDREITNILESTPGATFLLESAAHSDVGINNLRNNTISSNVEFH 179
Db      121 ELQVIDINDHSPFLDKQMLVKVSESSPGTFFPLKNAEDDVGQNNIENYIISPNYSFR 180
Qy      180 INVHNGSGNANYSELYLDKVDREVPRLRLTGLDGGSPRRSTTLIRLIVLDINDV 239
Db      181 VLTFRRSQGRKPELVLDKALDREBEALRLTLALDGGSPRRSTAAVYIEVLDVNDNA 240
Qy      240 PEFVSLKVOVPENSPGSLVTVVSARDLDTGSGEIVAFVFAVTEERTLTFRINSYG 299
Db      241 PEFEPFVRVQISEDSPPGFLVVKVSAVDVDTGVNGEISLSPFOASEIGTFKINPLTG 300
Qy      300 NLHLKAEINYEALIQTYTLTQAKDGGISGKCTVVVAVHTDINDNPELLMSLTSPIDEN 359
Db      301 EIELKKQLDFEKLQSYEVNIEARDAFTSGKCTVLIQYIDVNDHAPVETMSAFTSPIDEN 360
Qy      360 SPETVVAVFRIRDRDSSGNNAKAVCSIODHLPEVLKPSVENYTLTVERALDREKREYNI 419
Db      361 ABEIVVALEFSVSDLDGSENGKISCSIOEDLPELAK-SAENETTLTTERPLDRESAEYNI 419

```





70 ETIGELLNPLTGDLLNEKIDREELCGPTPECVLPFOLLLEKPFOTFRAELWVRDINDH 129

[illegible]

```

; INFORMATION FOR SEQ ID NO: 110
;
; SEQUENCE CHARACTERISTICS:
;
;   LENGTH: 787 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
US-08-268-161a-110

```

Query Match	71.3%;	Score 2287;	DB 1;	Length 787;
Best Local Similarity	72.3%;	Pred. No. 2.4e-207;		
Matches 447;	Conservative	67;	Mismatches 104;	Indels 0;
			Gaps	0;

QY	10	QKROVSLLCVFGVSMAGAEPLRFYVAEEERPGFLPANAIDGJGVEELSSARGCRISD	6
QY	10	QKROVSLLCVFGVSMAGAEPLRFYVAEEERPGFLPANAIDGJGVEELSSARGCRISD	6
Db	10	EQGQVILILLLLEVLTAQMEPRRYSVMEETBRGSAFVANAANDGJGVEELARGARAYSE	69
QY	70	ETIGFLILNPLTGDILLNEKIDREELCGPPEPCVLPOLLLEKPROIFRAELMYARDINDH	129
Db	70	DNBGQGLDQLQNGOILNEKIDREELCGPPEPCIMHGYVLLKPKPEVFAELVAYLDINDH	129
QY	130	SPVEFLDREITLLESTPGATFLLSAGHSADVGINNLRYTSSVWEFHINVDHNEGN	189
Db	130	SPFEPRREMTLKIPEPSSIGTVFPLKARDDVDGNNQVYNIPSHFVSTRGDDR	189
QY	190	VYSELVLDKVDREVEPVELRLTLTGLOGGSPRPSGTLIRIAYLDINDNVPEVSELYKV	249
Db	190	KIPELVLDRELDREDOAELRLTLAVDGGSPRSGTVQILILVLANDAPEFVQALYEV	249
QY	250	QVPENSPVGSLVTVVSARDLDTGSGNGEIVYAFYATERLTKTFRINSTGMLHKAELYN	309
Db	250	QVPENSPVGSLVTVVKSARDLDTGTNGELISYLSYSSOEDKPEFLSSLSGELRLKIKDF	309
QY	310	EALQITTLIOAKDGGGSGKCVVYVHTDININPELLMSSLTSPSPENSPYVAVFR	369
Db	310	EMSSSTVDDIEASDDGGGSGKCSVYVLDVNONPELISSTSPIDENSPREVEVALFR	369
QY	370	IRRODSGNNAKVCSTIOGHLFPVLPKSEVNFYLVATERALDRBERTYNTITYTDLGTP	429
Db	370	IRRODSGNGKMKCSTIODVFPKLPSEVNFYLVATERALDRBETRAEVNTITYTDLGTP	429
QY	430	RLKTQNLNVTVSDVNDNAPFEQSOTVYTLTVLRNNNSPALHISVSATRDSGANAQVYS	489
Db	430	RLKTEQISTVLVSDVNDNAPAFQTSYTLTVLRNNNSPALHISVSATRDSGTNAQVYS	489
QY	490	LILPHHPOLPLGSLVSTINADNGOLFALRSIDFALDAFEFRGADRGSPALSSQALYRV	549
Db	490	LILPQOPHLPLTSLVSTINTDNGHLFEALQSDYDIALDAFEFRVGAIDRGSPALSSQALYRV	549
QY	550	LVAADNAPAEVLYPLONGSAPCTELVPRAAEGYVJVAEVVAVDDSGQNMALSYOLIKA	609
Db	550	LVLADANDNSPEVLYPLONGSAPCTELVPRAAEGYLVTKVAVAVDDSGQNMALSYOLIKA	609
QY	610	TEBGLGVWAHNGEVYRTA	627
Db	610	TEBGLGVWAHNGEVYRTA	627

RESULT 3  
US-08-453-702A-110  
; Sequence 110, Application US/08453702A  
; Patent No. 5891706  
; GENERAL INFORMATION:  
; APPLICANT: Suzuki, Shintaro  
; TITLE OF INVENTION: Protocadherin Materials and Methods  
; NUMBER OF SEQUENCES: 115  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &  
; ADDRESSEE: Borum  
; STREET: 233 South Wacker, 6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606



Db 370 IRDRSGNGKMCISIODVFPKLPSEVNFYRLVTEGALDRERAEYNITITITDGT 429  
QY 430 RLKTOHNLTVVSDVNDNAPFEFOTTYTLRYRENNSPALHIGVSATDRSGANAOVTYS 489  
Db 430 RLKTEGOSTIVLSDVNDNAPFOTSTYTLRYRENNSPALHIGVSATDRSGANAOVTYS 489  
QY 490 LLPHPDPLPGSLVSIINADNGQLFALRSLDFEALQAFEEFVGAADRGSPALSSQALVRY 549  
Db 490 LLPHPDPLPGSLVSIINADNGHLFALQSLDYELQAFEEFVGAADRGSPALSSQALVRY 549  
QY 550 LVADANDAPFVLPYPLONGSAPCTELVPRAAEAGYLAKVAVDGDGSONAMLSYOLLA 609  
Db 550 LVADANDNSPFLVPLONGSAPCTELVPRAAEAGYLAKVAVDGDGSONAMLSYOLLA 609  
QY 610 TEPGLFGVMAHNGEVRTA 627  
Db 610 TEPGLFGVMAHNGEVRTA 627

## RESULT 5

PCT-US95-08071-110  
; Sequence 110, Application PC/TUS9508071  
; GENERAL INFORMATION:  
; APPLICANT: Suzuki, Shintaro  
; TITLE OF INVENTION: Protocadherin Materials and Methods  
; NUMBER OF SEQUENCES: 115  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/08071  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/12588  
; FILING DATE: 23 DEC 1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/998,003  
; FILING DATE: 29 DEC 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Noland, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 32149  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 110:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 787 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US95-08071-110

Query Match 71.3%; Score 2287; DB 5; Length 787;  
Best local similarity 72.3%; Pred. No. 2,4e-207;  
Matches 447; Conservative 67; Mismatches 104; Indels 0; Gaps 0;  
QY 10 QKQVSLLEVLGVSWAGAPLRYVAETERTGLANLAIDLGLGVEELSSARGCRIYSD 69  
Db 10 EQRQVLLILLLEVLGAWEPFRYSYMEETRGSEFVANLANLDLGLGVELAERGARVYSE 69

QY 70 ETIGLELLPLGDLILNKLREDEICGTEPCVLPFOLLLEKPPQIRAEIWRVDINDH 129  
Db 70 DNEQGLDLDLQGLQLINKLREKCGTEPCIMHFPVLLKPLEVRAELLYVDINDH 129  
QY 130 SPVFLDREITLNLISSTPGATFELLESADSDVGINLNLNYISSNVFHIIVHNGEGN 189  
Db 130 SPEFPEREMTLKIPETSLIGVFPPLKKAADLDVGSNNQVNIYSNHFHSTRRGGR 189  
QY 190 VSELYLVKVDREYPELRLITGLDGSPPRSGTTLRIIVLDINDNVEPEVSLYK 249  
Db 190 KYPELVLDLREBOAELRLITLAVDGSPPRSGTVOILILIVDANDNAEVEVQALVEY 249  
QY 250 QVPENSPVSLVTVSARDLDTGSNGEIVYAFVTERLTFRINSTGNHLKAEIYN 309  
Db 250 QVPENSPVSLVTVKSARDLDTGTNGELISLYSSQELDKPFELSSIGELRIKKLDF 309  
QY 310 EAIQTYTLITQAKDGGSLGKCTVYVHTDINDNPELLMSLSPIPENSPEVAVFR 369  
Db 310 ETMSYDLDIEASDGGSLGKCSVSKVLDVNDNPELISLSTSPIPENSPETVALFR 369  
QY 370 IRDRSGNNAKVCSTODHLPVULKPSVENFYTTLTERALDREKTEYNITITDGT 429  
Db 370 IRDRSGNGKMCISIODVFPKLPSEVNFYRLVTEGALDRERAEYNITITDGT 429  
QY 430 RLKTOHNLTVVSDVNDNAPFEFOTTYTLRYRENNSPALHIGVSATDRSGANAOVTYS 489  
Db 430 RLKTEGOSTIVLSDVNDNAPFOTSTYTLRYRENNSPALHIGVSATDRSGANAOVTYS 489  
QY 490 LLPHPDPLPGSLVSIINADNGQLFALRSLDFEALQAFEEFVGAADRGSPALSSQALVRY 549  
Db 490 LLPHPDPLPGSLVSIINADNGHLFALQSLDYELQAFEEFVGAADRGSPALSSQALVRY 549  
QY 550 LVADANDAPFVLPYPLONGSAPCTELVPRAAEAGYLAKVAVDGDGSONAMLSYOLLA 609  
Db 550 LVADANDNSPFLVPLONGSAPCTELVPRAAEAGYLAKVAVDGDGSONAMLSYOLLA 609  
QY 610 TEPGLFGVMAHNGEVRTA 627  
Db 610 TEPGLFGVMAHNGEVRTA 627

## RESULT 6

US-08-453-695A-112  
; Sequence 112, Application US/08453695A  
; Patent No. 5708143  
; GENERAL INFORMATION:  
; APPLICANT: Suzuki, Shintaro  
; TITLE OF INVENTION: Protocadherin Materials and Methods  
; NUMBER OF SEQUENCES: 115  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &  
; STREET: 233 South Wacker, 6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/453,695A  
; FILING DATE:  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5708143and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 32658  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448  
 TELEX: 25-3856  
 INFORMATION FOR SEQ ID NO: 112:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 797 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-453-695A-112

Query Match 62.6%; Score 2009.5; DB 1; Length 797;  
 Best Local Similarity 62.7%; Pred. No. 4.5e-181;  
 Matches 393; Conservative 83; Mismatches 148; Indels 3; Gaps 3;

QY 1 MEARVHALQKROVSLICFLGVSM-AGAEPLRFVAEETERTGLANLAIDLGVEEL 59  
 DB 1 METALAKIPQOROVFFLTI-LSLKMSSSEAIRYSMEETESGYVANAIAKLDGIRVCEL 59  
 QY 60 SARCRIVSDETIIGFLNPLTGDLINKEDREELGTEPCVLPFOLLLEKPFQIFRA 119  
 DB 60 SSGAQIHKKKKNKELLQDAETGNLFKRLKDRRELCGTEPCVLPFOLLLEKPFQIFOT 119  
 QY 120 ELWVRDINDHSPVFLDREITNILESTPGATFLLESADSDVGINNERNYSSNYFH 179  
 DB 120 ELQITDINDHSPFPKMKMLTIPESAHPGVFPKAAKRDSDIGSNAYQNTVNNLPHF 179  
 QY 180 INVHNGEGNYSSELYDKVDLDRREVEPLRLTLTGDLGSPRSRGTLLIRILVINDV 239  
 DB 180 VYHSRTDKRPVELVDRLDREDEPRLTLTLTGDLGAPRSRGTTVHIEVVIDINDS 239  
 QY 240 PEFVESLIRKVOYVENSVPVGLVYVSARDLDTGSGEIVYAFYATERTLKTFRINSTG 299  
 DB 240 PQVOSLIRKVOYVENNPLNAFVYVSATDLDAVGYNVYSLFQG-YGVFPVIDEITG 298  
 QY 300 NHTLKAELNVEAIQYITLTIOAKDGGSLGKCTVVVHTDINDPELLMSLTSPFIPEN 359  
 DB 299 EHLSELDEFEISNHNIEIATDGGSLGKCTVAVQVLDVNDNAPELLIRKTLVLPEN 358  
 QY 360 SPETVAVAFRIKRDGSGNNAKWCSTQDLHPVLPKPSVENFTLVTERLADREERTYNI 419  
 DB 359 SAETVAVAFVSVDSDGNGRWCSIPNNIPFLKPTFENYITLVTEGLDENREAYNI 418  
 QY 420 TITVVDLGTPLRKTQHNLTIVYSDVNDNAPTESQTYTLTVRENNSPALHIGSVATDRD 479  
 DB 419 TITVVDLGTPLRKTQHTTIVYSDVNDNAPATQTSYTFVHEHNSPALHIGISATDSD 478  
 QY 480 SGANQVYTSILPRHPDPLGSLVSNADNGQLFALRSLDFEALQAEFRVGAADRGSP 539  
 DB 479 SGNNAHTYSLPRDPPQLADSLISINVDNGQLFALRALDYEALQSEFYVGAADRGSP 538  
 QY 540 ALSSQALVRVLYADVANDNAPVLYPLONGSAPCTELVPRAAEGYLVAKVAVADGSDGN 599  
 DB 539 ALSQTLVRVLYDDNDNAPVLYPLONGSAPCTELLPRAAEPGYLITVAVAVDRDSDGN 598  
 QY 600 AWLSYOLLKATEPGLFGVAHNGEVRT 626  
 DB 599 AWLSYOLLKATEPGLFGVAHNGEVRT 625

RESULT 7  
 US-08-268-161A-112  
 Sequence 112, Application US/08268161A  
 Patent No. 5798224  
 GENERAL INFORMATION:  
 APPLICANT: Suzuki, Shintaro  
 TITLE OF INVENTION: Procoadherin Materials and Methods  
 NUMBER OF SEQUENCES: 115  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &  
 STREET: 233 South Wacker, 6300 Sears Tower  
 CITY: Chicago  
 STATE: Illinois

COUNTRY: USA  
 ZIP: 60606  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/268,161A  
 FILING DATE: June 27, 1994  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Young J. Suh  
 REGISTRATION NUMBER: P-41,337  
 REFERENCE/DOCKET NUMBER: 27866/32149  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 312/474-6300  
 TELEFAX: 312/474-0448  
 TELEX: 25-3856  
 INFORMATION FOR SEQ ID NO: 112:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 797 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-268-161A-112

Query Match 62.6%; Score 2009.5; DB 1; Length 797;  
 Best Local Similarity 62.7%; Pred. No. 4.5e-181;  
 Matches 393; Conservative 83; Mismatches 148; Indels 3; Gaps 3;

QY 1 MEARVHALQKROVSLICFLGVSM-AGAEPLRFVAEETERTGLANLAIDLGVEEL 59  
 DB 1 METALAKIPQOROVFFLTI-LSLKMSSSEAIRYSMEETESGYVANAIAKLDGIRVCEL 59  
 QY 60 SARCRIVSDETIIGFLNPLTGDLINKEDREELGTEPCVLPFOLLLEKPFQIFRA 119  
 DB 60 SSGAQIHKKKKNKELLQDAETGNLFKRLKDRRELCGTEPCVLPFOLLLEKPFQIFOT 119  
 QY 120 ELWVRDINDHSPVFLDREITNILESTPGATFLLESADSDVGINNERNYSSNYFH 179  
 DB 120 ELQITDINDHSPFPKMKMLTIPESAHPGVFPKAAKRDSDIGSNAYQNTVNNLPHF 179  
 QY 180 INVHNGEGNYSSELYDKVDLDRREVEPLRLTLTGDLGSPRSRGTLLIRILVINDV 239  
 DB 180 VYHSRTDKRPVELVDRLDREDEPRLTLTLTGDLGAPRSRGTTVHIEVVIDINDS 239  
 QY 240 PEFVESLIRKVOYVENSVPVGLVYVSARDLDTGSGEIVYAFYATERTLKTFRINSTG 299  
 DB 240 PQVOSLIRKVOYVENNPLNAFVYVSATDLDAVGYNVYSLFQG-YGVFPVIDEITG 298  
 QY 300 NHTLKAELNVEAIQYITLTIOAKDGGSLGKCTVVVHTDINDPELLMSLTSPFIPEN 359  
 DB 299 EHLSELDEFEISNHNIEIATDGGSLGKCTVAVQVLDVNDNAPELLIRKTLVLPEN 358  
 QY 420 TITVVDLGTPLRKTQHNLTIVYSDVNDNAPTESQTYTLTVRENNSPALHIGSVATDRD 479  
 DB 419 TITVVDLGTPLRKTQHTTIVYSDVNDNAPATQTSYTFVHEHNSPALHIGISATDSD 478  
 QY 480 SGANQVYTSILPRHPDPLGSLVSNADNGQLFALRSLDFEALQAEFRVGAADRGSP 539  
 DB 479 SGNNAHTYSLPRDPPQLADSLISINVDNGQLFALRALDYEALQSEFYVGAADRGSP 538  
 QY 540 ALSSQALVRVLYADVANDNAPVLYPLONGSAPCTELVPRAAEGYLVAKVAVADGSDGN 599  
 DB 539 ALSQTLVRVLYDDNDNAPVLYPLONGSAPCTELLPRAAEPGYLITVAVAVDRDSDGN 598  
 QY 600 AWLSYOLLKATEPGLFGVAHNGEVRT 626  
 DB 599 AWLSYOLLKATEPGLFGVAHNGEVRT 625

Db 599 AWLSFOLKATEPGLFSVAHNGEVRT 625

RESULT 8  
US-08-453-702A-112  
Sequence 112 Application US/08453702A  
Patent No. 5891706  
GENERAL INFORMATION:  
APPLICANT: Suzuki, Shintaro  
TITLE OF INVENTION: Protocadherin Materials and Methods  
NUMBER OF SEQUENCES: 115  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &  
ADDRESSEE: Borun  
STREET: 233 South Wacker, 6300 Sears Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/453,702A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5891706and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 32657  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 112:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 797 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-453-702A-112

Query Match 62.6%; Score 2009.5; DB 2; Length 797;  
Best Local Similarity 62.7%; Pred. No. 4.5e-181;  
Matches 393; Conservative 83; Mismatches 148; Indels 3; Gaps 3;

Db 1 MEARVHALOKROYSLLCVFLGVSW-AGAEPRLRYFAETERTGFLNLAIDGLGYEEL 59  
Db 1 METALAKIPOOROVFFLTI-LSLMKSSSEAIRSMPEETESGYMANLAKDGIKRGEL 59  
QY 60 SARCGRIYSDETIGFLLNPLTGLDLNKKIDREELCGPTEPCVLPOLLERKFOIFRA 119  
Db 60 SSRGAQIHYKNGKELLQDAETGNLFLKELDKRELKCGTEPEPCVLPOLLERKFOIFRA 119  
QY 120 ELWVRDINDHSPVLDREITNLIESTTPGATFLIESAHSDVGINNLKNTTSSNVYFH 179  
Db 120 ELQFLDINDHSPVLPNNLTPESAHPTGVPPLKARSDIGSNVQYTNPNLHFI 179  
QY 180 INVADNGEGNYSSELVDKYLDEEVEPELRLTGLDGGSPRSGTLLRIILVDINDY 239  
Db 180 VVTHSRDGRKPYELVDRLADREOPRELTLILALDGGAPSGTITVIEVVDINDS 239  
QY 240 PEYVESLKYOVSPNSVGLVYTVSARDIDTGSNGEIVYAFYATERTLKTFRINSYSG 299  
Db 240 POFVOSLKYOVSPNNLNFVTVSATIDAGYGNVTVSLPOG-YGVQRPVIDEITG 298  
QY 300 NLHKAELNEALOTYTLITQADGGGLSGKCTVYVAVTIDINNPPELMSSITSPEN 359  
Db 299 EHLKSKELDEFEISNHNIEIATDGGGLSGKCTVAVQVLDVNDNAPLELTKLTVLAVPEN 358  
QY 360 SPETVAVFRIIRDSDGNNAKWCISQDHLFVLKPSVENFYTLYVERALDREEREYNI 419

Db 359 SAETVAVFVSVDSDSGDNRCBWCISIPNNIPFLKTFENYITLYEGDLNRENEYNI 418  
QY 420 TITVDTLGPRLKTOHNLVTVSDVDNAPTISQTYTLIRVBNNSPALHISVSATDSD 479  
Db 419 TITVSDLGTPRLTOHNLVTVSDVDNAPTISQTYTLIRVBNNSPALHISVSATDSD 478  
QY 480 SGANAQVYSLPDPHPOPLGSLVSIINDNGOLFALRSIDPEALQAFEFVAGADRGSP 539  
Db 479 SGANAHTTSLPDPHPOPLGSLVSIINDNGOLFALRSIDPEALQAFEFVAGADRGSP 538  
QY 540 ALSSQALVRYLVADANDNAPVLYPLQNGSAPCTELVPRAAGVYLAQVAVDSDGON 599  
Db 539 ALSSQTLVRYLVADANDNAPVLYPLQNGSAPCTELVPRAAGVYLAQVAVDSDGON 598  
QY 600 AWLSYOLKATEPGLFSVAHNGEVRT 626  
Db 599 AWLSFOLKATEPGLFSVAHNGEVRT 625

RESULT 9  
US-09-099-639-112  
Sequence 112 Application US/09099639  
Patent No. 6262237  
GENERAL INFORMATION:  
APPLICANT: Suzuki, Shintaro  
TITLE OF INVENTION: Protocadherin Materials and Methods  
NUMBER OF SEQUENCES: 115  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &  
ADDRESSEE: Borun  
STREET: 233 South Wacker, 6300 Sears Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/099,639  
FILING DATE: 18 JUN 1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/263,161  
FILING DATE: 27 JUN 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Greta E. No. 6262237and  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/34703  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 112:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 797 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-099-639-112

Query Match 62.6%; Score 2009.5; DB 4; Length 797;  
Best Local Similarity 62.7%; Pred. No. 4.5e-181;  
Matches 393; Conservative 83; Mismatches 148; Indels 3; Gaps 3;

Db 1 MEARVHALOKROYSLLCVFLGVSW-AGAEPRLRYFAETERTGFLNLAIDGLGYEEL 59  
Db 1 METALAKIPOOROVFFLTI-LSLMKSSSEAIRSMPEETESGYMANLAKDGIKRGEL 59  
QY 60 SARCGRIYSDETIGFLLNPLTGLDLNKKIDREELCGPTEPCVLPOLLERKFOIFRA 119

```

Db 60 SSRGAQIHYKGNKELLQDAETGNLFLEKLDRELLCETPCVLIENPMQFOT 119
QY 120 ELWVNDINDHSPVFLDREITLTIESTPGATFELLESADSDVGINNRNTISSNVEFH 179
Db 120 ELQITDINDHSEFPNKKMLLTIPESAHPTVPLKARDSDIGSNAQNTVNPMLHFH 179
QY 180 INVHNGEGNYSSELYLKVLDREVEPELRLTGLDGGSPRSGTULIRLIVDINDV 239
Db 180 VVTHSRTOGRKPELVLDRLADREOPRELLITLALDGGAPSRSGTIVHLEVDINDNS 239
QY 240 PEFVESLYKQVNPENSPVGLVTVSARDLDTGSGEIVYAFVATERLTKTERINSTG 299
Db 240 PQFVQSLKYQVNPENNPINAFVTVSATDLGAGYGVNTVSLFQG-YGVFPFVIDEITG 298
QY 300 NHTLAELNVEALQTYTLTLOAKDGGGLSGCTVAVVHTDINDNPPELLMSLSPEN 359
Db 299 EHLKSKELDFEETISNHNIEATDGGISGCTVAVVQVLDVNDAPELTIRKLIYLVEN 358
QY 360 SPETVAVAFRIRDRDGSNNAKWCSIODHLPFLKPSVENEYTLVTERALDREERTENI 419
Db 359 SAETVAVAFVSVDSDSGDNGMVCISIPNNIPFLKPFENYTLVTEGPLDRENAEYNI 418
QY 420 TITVVDLGPRLKTOHNLTVTVSDVNDAPFTSOTTYTLRRENNSPALHIGSVATRD 479
Db 419 TTVSDLGTPRLTQHTTVQVSDINDNAPAFOTSTYMFVHNNSPALHIGTISATSD 478
QY 480 SCANAQVTVSLPDPHDPOLPLGSLVSINADNGOLFALRSLDEALQAEFRVGAADRS 539
Db 479 SCSNAHITYSLPDPDDPOLADSLISINVDNGOLFALRALDYEALQSEFEVYGATDGGSP 538
QY 540 ALSSQALVRLVADANDNAPVLYPLONGSAPCTELVPRAAEGLVAKVAVDGDGSON 599
Db 539 ALSSQTLVRLVADNDNAPVLYPLONASAPCTELPRAAEPGLITIKVAVDSDGON 598
QY 600 AWLSYOLKATEPGLFGVMAHNGEVRT 626
Db 599 AWLSFOLLKATEPGLFSVMAHNGEVRT 625

```

## RESULT 10

```

PCT-US95-08071-112
Sequence 112, Application PC/TUS9508071
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: Protocadherin Materials and Methods
NUMBER OF SEQUENCES: 115
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08071
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/12588
FILING DATE: 23 DEC 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/998,003
FILING DATE: 29 DEC 1992
ATTORNEY/AGENT INFORMATION:
NAME: Noland, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 32149

```

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ. ID NO: 112:
SEQUENCE CHARACTERISTICS:
LENGTH: 797 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-08071-112

```

```

Query Match 62.6%; Score 2009.5; DB 5; Length 797;
Best Local Similarity 62.7%; Pred. No. 4.5e-181;
Matches 393; Conservative 83; Mismatches 148; Indels 3; Gaps 3;

```

```

QY 1 MEARVVALQKROVSLICVFLGWSW-AGAEPLAFVFAVEERETFTLANLADIGLVEEL 59
Db 1 METALARIPOOROVFFLT-LSLIMKSSSEAIRISMEETESGYMANLAKDKIGIRGEL 59
QY 60 SARCGRIVSDETIGFLLNPLTGDLLNEKLDREELCGTPEPCVLPOLLLEKRFQIFRA 119
Db 60 SSRGAQIHYKGNKELLQDAETGNLFLEKLDRELLCETPCVLIENPMQFOT 119
QY 120 ELWVNDINDHSPVFLDREITLTIESTPGATFELLESADSDVGINNRNTISSNVEFH 179
Db 120 ELQITDINDHSEFPNKKMLLTIPESAHPTVPLKARDSDIGSNAQNTVNPMLHFH 179
QY 180 INVHNGEGNYSSELYLKVLDREVEPELRLTGLDGGSPRSGTULIRLIVDINDV 239
Db 180 VVTHSRTOGRKPELVLDRLADREOPRELLITLALDGGAPSRSGTIVHLEVDINDNS 239
QY 240 PEFVESLYKQVNPENSPVGLVTVSARDLDTGSGEIVYAFVATERLTKTERINSTG 299
Db 240 PQFVQSLKYQVNPENNPINAFVTVSATDLGAGYGVNTVSLFQG-YGVFPFVIDEITG 298
QY 300 NHTLAELNVEALQTYTLTLOAKDGGGLSGCTVAVVHTDINDNPPELLMSLSPEN 359
Db 299 EHLKSKELDFEETISNHNIEATDGGISGCTVAVVQVLDVNDAPELTIRKLIYLVEN 358
QY 360 SPETVAVAFRIRDRDGSNNAKWCSIODHLPFLKPSVENEYTLVTERALDREERTENI 419
Db 359 SAETVAVAFVSVDSDSGDNGMVCISIPNNIPFLKPFENYTLVTEGPLDRENAEYNI 418
QY 420 TITVVDLGPRLKTOHNLTVTVSDVNDAPFTSOTTYTLRRENNSPALHIGSVATRD 479
Db 419 TTVSDLGTPRLTQHTTVQVSDINDNAPAFOTSTYMFVHNNSPALHIGTISATSD 478
QY 480 SCANAQVTVSLPDPHDPOLPLGSLVSINADNGOLFALRSLDEALQAEFRVGAADRS 539
Db 479 SCSNAHITYSLPDPDDPOLADSLISINVDNGOLFALRALDYEALQSEFEVYGATDGGSP 538
QY 540 ALSSQALVRLVADANDNAPVLYPLONGSAPCTELVPRAAEGLVAKVAVDGDGSON 599
Db 539 ALSSQTLVRLVADNDNAPVLYPLONASAPCTELPRAAEPGLITIKVAVDSDGON 598
QY 600 AWLSYOLKATEPGLFGVMAHNGEVRT 626
Db 599 AWLSFOLLKATEPGLFSVMAHNGEVRT 625

```

## RESULT 11

```

US-07-998-003A-107
Sequence 107, Application US/07998003A
Patent No. 5643781
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: Protocadherin Materials and Methods
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
STREET: 20 South Clark Street

```

```

CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/998,003A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 564378land, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 30903
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/246-5750
TELEFAX: 312/984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 682 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-998-003A-107

Query Match          37.9%; Score 1217; DB 1; Length 682;
Best Local Similarity 41.0%; Pred. No. 3.3e-106;
Matches 254; Conservative 117; Mismatches 242; Indels 6; Gaps 4;

QY 12 ROVSLLCVFLGVSNAAGAPLRYFAEETERTGFTLANLAIDLGVEELSARGCRIVSDT 71
D 15 RVGVL-LLLGALKASTVHYEIPEREKGFAGVANNVANGDLGSLARRFPVSGAS 73
QY 72 IGFLLNPLTGDLINKREELCGTEPCVLPFQILLEPQIFRAELMVRDINDHSP 131
D 74 RRFVEVNETGEMVNDRLDEELCGTLPSCVTLELVENPLEFSVEVYIODINDNP 133
QY 132 VFDRREITNLLESTTGATFLESADSDVGINNLNRTISSNVYEHINVDNGECNVY 191
D 134 APFTQEMKLEISAVAGTRPLESHADPDGNSLQTYELSNVEFALRVOTREDSTKY 193
QY 192 SELVDKVLDEEYVPELRLTLTGDDGSPRSSTTLRLVLVDINDVPEVESLYKVQY 251
D 194 AELVLERALDEREPISQVLVTALDGTPLASLPIHIKVIDANDNAVFNOSLYRARV 253
QY 252 PENSPPGSLVVTASARDLDTSGNGEIVYAF-FYATERLTLPFRINSTGNLHKAEINYE 310
D 254 PGCGTSGTRVQVLAITDDEGPNGEIITISFGSHNRAGVRQLPALDVTGLTKGRDPE 313
QY 311 AIQTYTLTQAKDGG--LSGCTVVVHYVDINDNPPELLMSSITSPENSPETVAVAF 368
D 314 DTKLHEIYIOAKDKGANPEGACHKVLVEVDVNDNAPEITVSYSVPEDEASGVYIAL 373
QY 369 RIRDRSGNNAKMVCISIODHLPVLRKPSVENFTTLVTERALDREERTETVITVDIGT 428
D 374 SVTDLAGENGVLTCVPPGLPESLISLKNFTLTKTSADLDREYVPEVNSTIARADGT 433
QY 429 PRKTOHNTLVVSDVNDNAPTFSQTTTTLVRRENNSPALHIGSVATDRDGSANNAVY 488
D 434 PSLSALTITRYOVSDINDNPPOSSOSIYVIEENNLPAPIILNLSVWDPDAPQNRISF 493
QY 489 SLIPDPDPLIGSLVSNADNGQLFALRSLDEALQAEFRVGAADRSPALSSQALVR 548
D 494 FLDEQAGELGVGRYFTINRNGIYVSLVPLDYEDRREBELFAHISDGTVPYLAINISN 553
QY 549 VIVADANDNAPFLYLPLOGSAPCELTLYVRAEAGTLVAKVAVVGDSDQANMLSTYOLIK 608
D 554 IFVTDNRNAPVLYLPRPGSS--VEMLPKGTSGAHLVSVVGVWDADAGHNAMLSYSLFG 611

```

```

QY 609 ATEPGLFGVMAHNGEVETA 627
D 612 SPNQSLEFALGHTGQISTA 630

RESULT 12
US-08-453-274B-107
Sequence 107, Application US/08453274B
Patent No. 5663300
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: Protocadherin Materials and Methods
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,274B
FILING DATE: 30-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: No. 5663300and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 32660
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 682 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-453-274B-107

Query Match          37.9%; Score 1217; DB 1; Length 682;
Best Local Similarity 41.0%; Pred. No. 3.3e-106;
Matches 254; Conservative 117; Mismatches 242; Indels 6; Gaps 4;

QY 12 ROVSLLCVFLGVSNAAGAPLRYFAEETERTGFTLANLAIDLGVEELSARGCRIVSDT 71
D 15 RVGVL-LLLGALKASTVHYEIPEREKGFAGVANNVANGDLGSLARRFPVSGAS 73
QY 72 IGFLLNPLTGDLINKREELCGTEPCVLPFQILLEPQIFRAELMVRDINDHSP 131
D 74 RRFVEVNETGEMVNDRLDEELCGTLPSCVTLELVENPLEFSVEVYIODINDNP 133
QY 132 VFDRREITNLLESTTGATFLESADSDVGINNLNRTISSNVYEHINVDNGECNVY 191
D 134 APFTQEMKLEISAVAGTRPLESHADPDGNSLQTYELSNVEFALRVOTREDSTKY 193
QY 192 SELVDKVLDEEYVPELRLTLTGDDGSPRSSTTLRLVLVDINDVPEVESLYKVQY 251
D 194 AELVLERALDEREPISQVLVTALDGTPLASLPIHIKVIDANDNAVFNOSLYRARV 253
QY 252 PENSPPGSLVVTASARDLDTSGNGEIVYAF-FYATERLTLPFRINSTGNLHKAEINYE 310
D 254 PGCGTSGTRVQVLAITDDEGPNGEIITISFGSHNRAGVRQLPALDVTGLTKGRDPE 313
QY 311 AIQTYTLTQAKDGG--LSGCTVVVHYVDINDNPPELLMSSITSPENSPETVAVAF 368
D 314 DTKLHEIYIOAKDKGANPEGACHKVLVEVDVNDNAPEITVSYSVPEDEASGVYIAL 373
QY 369 RIRDRSGNNAKMVCISIODHLPVLRKPSVENFTTLVTERALDREERTETVITVDIGT 428

```



```

Db 374 SVTDLGAGENGLVCEVPPGLPFLSTSLKNTYFLKTSADLRETYEVNLSIARDAGT 433
QY 429 PRKTOHNLTVTVSDVNDNAAPFSSQTYTLRVRENNSPALHIGSVKATDDSGANAOVTY 488
Db 434 PSLALTIVRVQVSDINDNPQSSQSSDYVIEENNLPGAPILNLSWDDPADPQNALSLF 493
QY 489 SLPPHDPOLPLGSLVSVINADNGQLFALRSIDFALQAFEFVGAAGRSGSPALSSQALVR 548
Db 494 FLLEOGAETGLVGRYFTINRNGIVSSLPDYDRREFELTAHISDGTPLVATNLSVN 553
QY 549 VLVADANDANAPFLVPLQNGSAPCTELVPRAAEAGYLVAKVAVVGDGSONAMLSTYOLK 608
Db 554 IFVTDNRDANAPQVLYXPRGGSS--VEMLPRTGSAGHLVSRVVGWDADAGHNAWLSTSLFG 611
QY 609 ATEPGLGVMAHNGEVRTA 627
Db 612 SPNOSLFAIGLHTGOISTA 630

```

## RESULT 13

```

US-08-453-695A-107
; Sequence 107, Application US/08453695A
; Patent No. 5708143
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: Protocadherin Materials and Methods
; NUMBER OF SEQUENCES: 115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
; STREET: 233 South Wacker, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,695A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5708143and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 32658
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 682 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-453-695A-107

```

Query Match 37.9%, Score 1217, DB 1, Length 682;

Best Local Similarity 41.0%, Pred. No. 3,3e-106;

Matches 254; Conservative 117; Mismatches 242; Indels 6; Gaps 4;

```

QY 12 ROVSLICVFLGVSWAGAPLRYFAVEBERGTFLLANLIDLGVEELISANGRIYSDET 71
Db 15 RVYGVLT-LITGALINKASVTIHEIPEERKGFAGVAVVANGLDIGLSISARFPVSGAS 73
QY 72 IGFLLNPLTGLDLNLEKLDREELGPRPCVLPFOLLLEKPFQFERAELAVRDINDHSP 131
Db 74 RREFVNRBTGEMFVNDRLREELGLTLPSCVTILELVENPLELSEVAVVIOIDINDNP 133

```

```

QY 132 VFLEDRITLILESTPGATFELLESADSDVGINNLRNTYSSNVYEHINVDHNGEVNY 191
Db 134 APFLOEMKLEISEAVAPGTRFPLESAHDPLDLSNSIQVTELSRNEYFAIRVQTRDSTKY 193
QY 192 SELVLDKVDREVEPELRKLTITGDOGSPRPGSTTLIRIVLDINDNPEFESLYXQV 251
Db 194 AELVLERALDREREPSLQVLTALDGTLPALSASLPRIKIVLDANDNAVEFQSLYRKY 253
QY 252 PENSPIVSLVTVVSARDLDTGSGEIVYAF-FYATERLTKEPRINSTSGNLKALAYE 310
Db 254 PGCGTGTIRVOYLALDDEGRNGEITTFSGSHNAGVQALFADLYTGMLTKRDLPE 313
QY 311 AIQVTLTLTQAKDGG--LSGCTVVVHTDINDNPELLMSLSPSPENSPEVYAVF 368
Db 314 DFKLHEIYIQAKDKANPGACNKVLEVEVDVNDNAPEITVSVSPVEDASGVYIAL 373
QY 369 RIRDRSGNNAKAVGSIQDHLPLVLRKPSVENRTYTLTEALDREREVTNITITDGT 428
Db 374 SVTDLGAGENGLVCEVPPGLPFLSTSLKNTYFLKTSADLRETYEVNLSIARDAGT 433
QY 429 PRKTOHNLTVTVSDVNDNAAPFSSQTYTLRVRENNSPALHIGSVKATDDSGANAOVTY 488
Db 434 PSLALTIVRVQVSDINDNPQSSQSSDYVIEENNLPGAPILNLSWDDPADPQNALSLF 493
QY 489 SLPPHDPOLPLGSLVSVINADNGQLFALRSIDFALQAFEFVGAAGRSGSPALSSQALVR 548
Db 494 FLLEOGAETGLVGRYFTINRNGIVSSLPDYDRREFELTAHISDGTPLVATNLSVN 553
QY 549 VLVADANDANAPFLVPLQNGSAPCTELVPRAAEAGYLVAKVAVVGDGSONAMLSTYOLK 608
Db 554 IFVTDNRDANAPQVLYXPRGGSS--VEMLPRTGSAGHLVSRVVGWDADAGHNAWLSTSLFG 611
QY 609 ATEPGLGVMAHNGEVRTA 627
Db 612 SPNOSLFAIGLHTGOISTA 630

```

## RESULT 14

```

US-08-268-161A-107
; Sequence 107, Application US/08268161A
; Patent No. 5798224
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: Protocadherin Materials and Methods
; NUMBER OF SEQUENCES: 115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
; STREET: 233 South Wacker, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/268,161A
; FILING DATE: June 27, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Young J. Suh
; REGISTRATION NUMBER: P-41,337
; REFERENCE/DOCKET NUMBER: 27866/32149
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 682 amino acids

```

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
us-08-268-161a-107

Query Match 37.9%; Score 1217; DB 1; Length 682;  
Best Local Similarity 41.0%; Pred. No. 3,3e-106;  
Matches 254; Conservative 117; Mismatches 242; Indels 6; Gaps 4;

```

QY 12 ROYSLICVFLGYSWAGAEPLRYFAEETERTGTFELANLAIDGLVEELSGRCRIVSDT 71
DB 15 RVYGVV-LTLGALNKASVYIHEIPEERKGFAGVNVANGLDGLGSLASARFPVVSAS 73
QY 72 IGFLLNPLTGLDLINEKIDREELCGPTEPCVLPOLLERFQIFRAELVWRDINDHSP 131
DB 74 RREFEVNETGEMFVNDRLDRRELGLTLPSCVTLELVENPLELFSEVYIODINDNNP 133
QY 132 VFLEDTLNLLESTTPGATFELLESADSDVGINNLRYNTSSNYFHIHNDGEGNVY 191
DB 134 APTQEMKLEISEAVAPGTRFPLESAHDPDGLNSLQTELSRNEYFALRVQTRDSTKY 193
QY 192 SELVLDKVLDRREVELRLTTLTGDDGSPRRSGTTLRLILVLDINDVPEVESLYKVQV 251
DB 194 AELVLERALDRERPSLQVLTLALDGGTPALASLPIHIKVLANDANAPVFNQSLYRVRV 253
QY 252 PENSFVGLVYTVSARDLDTGSGEIVYAF-FYATERLTKFRINSGNLHKAELNVE 310
DB 254 PGGCTSGTRVQVATLTDDEGPNGEIITSPGSHRAGVROLFAIDLVTGMITIKGRDLE 313
QY 311 AIQYITLTIOAKDGGG--LSGKCTVVVHTDINDNPPELLMSLTPSPENSPETVAVF 368
DB 314 DTKLHEIYIOAKDKGANEGAHCKVLEVVVNDNNAPEITVTSYSPVEDASGTVAL 373
QY 369 RIRPDSGNNAKWCISIODHLPEVLPKPSVENFYLTERALDREREYNTITVTDLGT 428
DB 374 SVTDLDAENGLVTCVPPGLPFSLSLKNYFTLKTSADLDRETYEYMLSTIARAGT 433
QY 429 PRKTOHNLTVYSDVNDNAPTESQTYTLRVRENNSPALHIGSVATDRDSCANAQVY 488
DB 434 PLSALITVRYOVSDINDNPQSSQSYDYIEENNLPGAPILNLSWMDADAPQANRLSF 493
QY 489 SLPPHPDPLPLGSLVSIINDNGQLFRLSLDEPALOAEFFRVGAADRGSPALSSQALVR 548
DB 494 FLEBQAGETGLVGRYFTINDNGIVSSLVPLDYEDRREFELTAHISDGTPTVATINSV 553
QY 549 VLVADANDNAPFVLYPLONGSAPCTELVPRAEAGYLVAKVAVVDDGSGONAMLSYQLK 608
DB 554 IFVTDNRDNDAPQVLYPRPGSS--VEMLPRTGSAGHLVSRVVGMDADAGHNAWLSTSLFG 611
QY 609 ATEPGLFGVAHNGEVRTA 627
DB 612 SPNOSLFAIGLHTGOISTA 630

```

## RESULT 15

US-08-453-702A-107  
Sequence 107, Application US/08453702A  
Patent No. 5891706

GENERAL INFORMATION:  
APPLICANT: Suzuki, Shintaro  
TITLE OF INVENTION: Protocadherin Materials and Methods  
NUMBER OF SEQUENCES: 115  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &  
ADDRESS: 233 South Wacker, 6300 Sears Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/453,702A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5891706and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 32657  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ. ID NO: 107:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 682 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
us-08-453-702A-107

Query Match 37.9%; Score 1217; DB 2; Length 682;  
Best Local Similarity 41.0%; Pred. No. 3,3e-106;  
Matches 254; Conservative 117; Mismatches 242; Indels 6; Gaps 4;

```

QY 12 ROYSLICVFLGYSWAGAEPLRYFAEETERTGTFELANLAIDGLVEELSGRCRIVSDT 71
DB 15 RVYGVV-LTLGALNKASVYIHEIPEERKGFAGVNVANGLDGLGSLASARFPVVSAS 73
QY 72 IGFLLNPLTGLDLINEKIDREELCGPTEPCVLPOLLERFQIFRAELVWRDINDHSP 131
DB 74 RREFEVNETGEMFVNDRLDRRELGLTLPSCVTLELVENPLELFSEVYIODINDNNP 133
QY 132 VFLEDTLNLLESTTPGATFELLESADSDVGINNLRYNTSSNYFHIHNDGEGNVY 191
DB 134 APTQEMKLEISEAVAPGTRFPLESAHDPDGLNSLQTELSRNEYFALRVQTRDSTKY 193
QY 192 SELVLDKVLDRREVELRLTTLTGDDGSPRRSGTTLRLILVLDINDVPEVESLYKVQV 251
DB 194 AELVLERALDRERPSLQVLTLALDGGTPALASLPIHIKVLANDANAPVFNQSLYRVRV 253
QY 252 PENSFVGLVYTVSARDLDTGSGEIVYAF-FYATERLTKFRINSGNLHKAELNVE 310
DB 254 PGGCTSGTRVQVATLTDDEGPNGEIITSPGSHRAGVROLFAIDLVTGMITIKGRDLE 313
QY 311 AIQYITLTIOAKDGGG--LSGKCTVVVHTDINDNPPELLMSLTPSPENSPETVAVF 368
DB 314 DTKLHEIYIOAKDKGANEGAHCKVLEVVVNDNNAPEITVTSYSPVEDASGTVAL 373
QY 369 RIRPDSGNNAKWCISIODHLPEVLPKPSVENFYLTERALDREREYNTITVTDLGT 428
DB 374 SVTDLDAENGLVTCVPPGLPFSLSLKNYFTLKTSADLDRETYEYMLSTIARAGT 433
QY 429 PRKTOHNLTVYSDVNDNAPTESQTYTLRVRENNSPALHIGSVATDRDSCANAQVY 488
DB 434 PLSALITVRYOVSDINDNPQSSQSYDYIEENNLPGAPILNLSWMDADAPQANRLSF 493
QY 489 SLPPHPDPLPLGSLVSIINDNGQLFRLSLDEPALOAEFFRVGAADRGSPALSSQALVR 548
DB 494 FLEBQAGETGLVGRYFTINDNGIVSSLVPLDYEDRREFELTAHISDGTPTVATINSV 553
QY 549 VLVADANDNAPFVLYPLONGSAPCTELVPRAEAGYLVAKVAVVDDGSGONAMLSYQLK 608
DB 554 IFVTDNRDNDAPQVLYPRPGSS--VEMLPRTGSAGHLVSRVVGMDADAGHNAWLSTSLFG 611
QY 609 ATEPGLFGVAHNGEVRTA 627
DB 612 SPNOSLFAIGLHTGOISTA 630

```

Search completed: May 16, 2003, 10:49:15  
Job time: 20 secs

Fri May 16 11:36:27 2003

us-10-050-704-125.ra1

Page 11





Db 420 ALRSLDYLALQAEFRVCAATDRGSPALSSSEALVRLVLDANDNSPFVLYPLONGSAPCTE 479  
 QY 575 LVPRAAEAGYLVAKVAVAVDGSQONAMLSYQLLKATEPGLGVMAHNGEVRTA 627  
 Db 480 LVPRAAEAGYLVTKVAVAVDGSQONAMLSYQLLKATEPGLGVMAHNGEVRTA 532

## RESULT 2

T09055  
 Protocadherin 68 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 11-Jun-1999  
 C:Accession: T09055  
 R:Jin, F.; Xu, H.; Israel, D.  
 Submitted to the EMBL Data Library, October 1997  
 A:Reference number: 216540  
 A:Accession: T09055  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-889 <JIN>  
 A:Cross-references: EMBL:AF029343; NID:g2599501; PID:g2599502  
 C:Genetics:  
 A:Gene: PCH68

Query Match 34.1% Score 1094; DB 2; Length 889;  
 Best Local Similarity 38.5% Pred. No. 1,3e-64;  
 Matches 252; Conservative 116; Mismatches 235; Indels 52; Gaps 16;

QY 14 VSLCLVFLGVSWAGA---EPLRYFAEETRGTFPLANLALDIDGLVEELSGARCVSDRTIGFLL-----NPLTG 82  
 Db 3 LSIICCCFL--LWAPALTLKLNLSVPEEGAGTYIGNIGBARQPLPAEKGCGGRSK 60  
 QY 64 ---CRIYSDETIGFLNPLTGLDLLNEKLDRELCPTEPCVLPFOLL--EKPFQIFRA 119  
 Db 61 SGSTRVLNSAPRHLDDVADSGLLYTKORIDRESLCRNNAKCOLSLHFVANDKEICMIKV 120  
 QY 120 ELAWNDIDHSPVFLDREITLNLSTTPGATFLESASHPDVGNNLRNYTSSNVY-- 177  
 Db 121 E--IQDINDNAPRSSQIYXIDISENAPGTRPPLTSAHDPDAGEGNLRYLLLRDQHL 178  
 QY 178 FHNVDHNGEGNVYSELVLDKVDREVPBELRLTLTGIDGSPRSSTGLIRILVLDIND 237  
 Db 179 FGLVYKSGDGDTKPEPLVYQKALDRQGNHHTLVLTALDGGEPFRSATVQINNVIDSND 238  
 QY 238 NVPEFVESLYKVQYPPSPGSLVYVSARDLGTSGNGEIVYAF--FYATERTLKTFRINS 296  
 Db 239 NSPVEAPSYLVELPENTPLGTVIDLNATDADGEPNGEVLVSFSYVPRVRELEFSIDP 298  
 QY 297 TSGMLHKAELNTEAIIQTYTLTIQAKDGG--LSGCTVAVVHTDINDNPEL-----LM 349  
 Db 299 KTGILIRKGNLDYENGMLEIDVQARLGLPRLIPAHCKVTVKILDRNDNAPSIGFVSVRO 358  
 QY 350 SLSLTPSPENSPETVAVAFRIIDRDSGNNAKWCIS-----QDHLFPVVK 394  
 Db 359 GALSSEAP---PGTIVIALVRYTRDSDGKNGQLCCRVLGGGGTGGGGGLGPGGSVFKE 415  
 QY 395 PSVNEETVYTERALDREERTENITTYVDLGTPLKTKHTNLTIVSYDNDNAPFSSQT 454  
 Db 416 ENIDNFTVTVDRPMDEYEDENVTIVADGSSPPLNSTKSPAILDLNDNDPREFTKG 475  
 QY 455 TYTLRVRNNSPALHISVSATDRDGSANQVYTSLLPPDPOLPGSLVSTINADGOLF 514  
 Db 476 LTVYQVHNNIPGDYXLSGLAOPDGLQNGTSTSLPSHIGVSTIYTVSVNPTGATY 535  
 QY 515 ALRSLDEALQAEFRVGAADRGSPA--LSSQALVRLVLDANDNAPFVLYP--LQNGSAPC 572  
 Db 536 ALRSENFEGTKAEFEKVLANDSGAPAHLESNATVRYTVLDVNNANVYIPLTIQNDTA-- 593  
 QY 573 TEL--VPRAAEAGYLVAKVAVAVDGSQONAMLSYQLLKATEPGLGVMAHNGEVRT 626  
 Db 594 -ELQVPRNAGIYLVSTVRLDSDDFGSGRLTEIYVDGNDHLEFIDPSSGEIRI 647

## RESULT 3

T31066  
 vascular cadherin-2 - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 21-Jul-2000  
 C:Accession: T31066  
 R:Tejo, P.; Breviarlo, F.; Huber, P.; Panzerl, C.; Dejana, E.  
 J. Biol. Chem. 273, 17565-17572, 1998  
 A:Title: Identification of a novel cadherin (vascular endothelial cadherin-2) located  
 A:Reference number: Z20962; MUID:98316522; PMID:9651350  
 A:Accession: T31066  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1180 <TEJ>  
 A:Cross-references: EMBL:F08715; NID:g2764993; PIDN:CAA69965.1; PID:g2764994  
 A:Experimental source: brain capillary

Query Match 28.1% Score 900.5; DB 2; Length 1180;  
 Best Local Similarity 34.7% Pred. No. 1.4e-51;  
 Matches 218; Conservative 114; Mismatches 241; Indels 55; Gaps 11;

QY 31 LRYFAEETRGTFPLANLALDIDGLVEELSGARCVSDRTIGFLL-----NPLTG 82  
 Db 32 VKFYTEEPVSGTVIGKLS-----QELRVERRGKAGDAFQILLQLPALPVQNSDGG 84  
 QY 83 DLLNEKLDRELCPTEPCVLPFOLLLEKPPQIFRAELWVDINDHSPVFLDREITLNI 142  
 Db 85 LSTSGRLDRKLCQDEDCVSPVLTAGASALHVEIQLVDINDHQPOFPKQDELEI 144  
 QY 143 LESTTGATFLESASHPDVGNNLRNYTSSNVYFHNVDHNGEGNVYSELVLDKVD 202  
 Db 145 SESASLHTRIPIDRLADDTGPNLSYSLSPSEHFALDYVGPDETHAEIYVVKELDR 204  
 QY 203 EEPBELRLTLTGIDGSPRSSTGLIRILVLDINDNVEPEFESLYKVQYPPSPGSLVY 262  
 Db 205 ELHSYFDVLTAYDNGNPPKSGISYKAVNVDSDNNSVFAESSIALIPEDYPTGILLI 264  
 QY 263 TVSARDLGTSGNGEIVYAF--FYATERTLKTFRINSTSGNLHKAELNTEAIIQTYTL 321  
 Db 265 NLTATDPDGPNGEVEFEFGKHVSPEVMNTFGIDAKQIILRLQLDYERKPAVEVDVQA 324  
 QY 322 KDGG--GLSGCTVAVVHTDINDNPEL--SLSLTPSPENSP--EYVAVAFRIIDRDSG 376  
 Db 325 RDGPNSTIPGHCKVLTKVDVNDNAPSLITWASQTSLSVSDLPDRSFSLVANSANDLSDG 384  
 QY 377 NNAKWCISIODHL--PVLKPSVENFYTLTERALDREERTENITTYVDLGTPLKTKH 435  
 Db 385 NNGLVHGMNLNGBELGHRKRTNGNTYMLTNATIDREQMPLYTLTVFQDGPPLSAEK 444  
 QY 436 NLTIVSYDNDNAPFSSQTYTLRVRNNSPALHISVSATDRDGSANQVYTSLLPPHD 495  
 Db 445 ELQIOWSDNDNAPVEFKSRSEVSTWENNPPSLHILITLKAHDADGSGKGYRIL----- 499  
 QY 496 POLPGSLVSTINADGOLFALRSIDFEALQAEFRVGAADRGSPALSSQALVRLVLDAN 555  
 Db 500 KDSVSHLVITDEFETGEVTAQRSIDYEQWAGEROVAEDRGQPOLASSISVWWSLIDAN 559  
 QY 556 DNAPFVLYP--LQNGSAPCTELV-----PRAEA--GYLVA 587  
 Db 560 DNAPFVLYP--LQNGSAPCTELV-----PRAEA--GYLVA 587  
 QY 588 KVVAVDGSQONAMLSYQLLKATEPGLF 615  
 Db 620 TIVARDADSGANGELFYSIQGNDHLEF 647

## RESULT 4

T00041  
 Bh-protocadherin PCDH7 (clone Bh-Pcdh-b) - human  
 C:Species: Homo sapiens (man)  
 C:Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 01-Dec-2000  
 C:Accession: T00041; T00040  
 R:Yoshida, K.; Yoshitomo-Nakagawa, K.; Seki, N.; Sasaki, M.; Sugano, S.

Genomics 49, 458-461, 1998  
 A>Title: Cloning, expression analysis, and chromosomal localization of BH-protocadherin  
 A:Reference number: 214074; MUID:98277460; PMID:9615233  
 A:Accession: T00041  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1072 <YOS>  
 A:Cross-references: EMBL:AB006756; NID:g2979419; PIDN:BAA25195.1; PID:g2979420  
 A:Experimental source: clone BH-Pcdh-b  
 A:Accession: T00040  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1058, 'M', '1060', 'LH', '1063', 'Y', '1065', 'TVFG' <YOS>  
 A:Cross-references: EMBL:AB006755; NID:g2979417; PIDN:BAA25194.1; PID:g2979418  
 A:Experimental source: clone BH-Pcdh-a  
 C:Genetics:  
 A:Map position: 4p15

Query Match 25.2%; Score 807; DB 2; Length 1072;  
 Best Local Similarity 30.4%; Pred. No. 2e-45;  
 Matches 207; Conservative 110; Mismatches 246; Indels 118; Gaps 15;

18 CVELGVSV--AGAPLRYFAETERTGTFLANLIDGL--GVEELARGCRIVSDETI 72  
 18 CILPLSFSLAAKOLLRLRLAEGBADVRIQNVASDLGIYTGSEVT-----ESLESGS 72  
 73 GFLINPLTGDLLE-KIDREEL--CG---PTEPCVLPFOLLLEKPRQ---IFRAELW 122  
 73 ELKIDNLGELSTERRIDREKLPQCCMIIDENECFLDEFESVIGPSQSWVDLEGGVY 132  
 123 VBDINDHSPFLDREITLILESTPGATFLLESADSDVGINLNKNTI----- 172  
 133 VLDINDNPTFSPVLTLLVEENRPVGTLLPTATDRDFGNHGERELLQEPGGGSG 192  
 173 -----SSNYFHIWHDNG 186  
 193 GSRAGADAPYPGGGNGASGGSGSKRLDASBEGGCTNPGCSVFEELQVADTP 252  
 187 EGNVSELYDLVDREVEPELRLTLTGDDGSPRSGTTLRLIVLINDNVPFVESL 246  
 253 DEKQPOLIVKALDREGDSYELFLRVADGDDPPRSSQAILRLVITVDNNSPFEEKSV 312  
 247 YKQVPEPNSPVGLVTVSARDLDTGSNGEIVYAFYATERTLKTFRINSTGNIHLKAE 306  
 313 YEADLAENSPAGTPILQLRAADLDVGVNGQIEYFGAATESVRLLRLDETSGWLVLHR 372  
 307 LMYEALQTYTLTQAKDG--GLSGKCYVYVHTDINDNPEL-----LMSLTSP 355  
 373 IDREVNOLRFTVMARDGQPPRTKATYVNLKIDENNVPSIEIRKIGRIPLKGVANV 432  
 356 IENSPETVAVFRIIDRDSGNNAKWCSTIODHLPFLKPSVE-----NFYTLVTERA 408  
 433 AEDVLTFTIALVQVSDRQENGVTCTVGVDPFOLKPADDTGDDQKKKRYFLHTSTP 492  
 409 LDRETEYNTITTYDICTPRLKTOHNLTVSVNDNAPFESQTYTLRYRENNSPAL 468  
 493 LYEATEFRENVIYAVDSGSPSLSSKNSLTVKGDTPNDPMEFGSVVEYYPENNIGE 552  
 469 HIGSVATCRDSCANAQVYISLPPHDPQLPLGSLVSTINADNGOLFALRSIDFELQAF 528  
 553 RYATVATDADSGKNAELAYSL-----DSSVMGFALIDPDGDLVTVVTDREOTDRE 606  
 529 FRYGADRSPPALSSQALRVLVADANDAP-----FVLYLQNGSAPCTELVRAEA 582  
 607 FVYNAKDKGIPVLOGSTVIYVADKNDNDPKFMQDVFIFYKEN-----LQPNSP-- 657  
 583 GYLAKVAVVDGSGONAWLS 603  
 658 ---VGMVTVMDADKGRNAEMS 675

RESULT 5  
 T00043

BH-Protocadherin-a - mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 05-Nov-1999  
 C:Accession: T00043  
 R:Yoshida, K.  
 Submitted to the EMBL Data Library, August 1997  
 A:Reference number: 214075  
 A:Accession: T00043  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1069 <YOS>  
 A:Cross-references: EMBL:AB006758; NID:d1227200; PIDN:BAA32597.1; PID:d1033562  
 C:Genetics:  
 A:Gene: pcdh7  
 A:Map position: 5C3-D

Query Match 24.9%; Score 799; DB 2; Length 1069;  
 Best Local Similarity 30.8%; Pred. No. 6.7e-45;  
 Matches 209; Conservative 105; Mismatches 247; Indels 118; Gaps 17;

17 LCVELGVSMGAEPRLRYFAETERTGTFLANLIDGL--GVEELARGCRIVSDETI 74  
 23 LCSTSLA---AAKOLLRLRLAEGBADVRIQNVASDLGIYTGSEVT-----ESLESGSEY 74  
 75 LLLNPLTGDLLE-KIDREEL--CG---PTEPCVLPFOLLLEKPRQ---IFRAELW 124  
 75 LKIDNLGELSTERRIDREKLPQCCMIIDENECFLDEFESVIGPSQSWVDLEGGVY 134  
 125 DINDHSPFLDREITLILESTPGATFLLESADSDVGINLNKNTI----- 172  
 135 DINDNPTFSPVLTLLVEENRPVGTLLPTATDRDFGNHGERELLQEPGGGSGE 194  
 173 -----SSNYFHIWHDNG 188  
 195 GRLGPDASAPYPGGGNGASGGSGSKRLDASBEGGCTSPSGSSVFEELQVADTP 254  
 189 NVSELYDLVDREVEPELRLTLTGDDGSPRSGTTLRLIVLINDNVPFVESLX 248  
 255 EKQPOLIVKALDREGDSYELFLRVADGDDPPRSSQAILRLVITVDNNSPFEEKSV 314  
 249 YQVPEPNSPVGLVTVSARDLDTGSNGEIVYAFYATERTLKTFRINSTGNIHLKAE 308  
 315 ADLAENSPAGTPILQLRAADLDVGVNGQIEYFGAATESVRLLRLDETSGWLVLHR 374  
 309 YEALQTYTLTQAKDG--GLSGKCYVYVHTDINDNPEL-----LMSLTSP 357  
 375 REEVNOLRFTVMARDGQPPRTKATYVNLKIDENNVPSIEIRKIGRIPLKGVANV 434  
 358 ENSPETVAVFRIIDRDSGNNAKWCSTIODHLPFLKPSVE-----NFYTLVTERALD 410  
 435 DVLDTFTIALVQVSDRQENGVTCTVGVDPFOLKPADDTGDDQKKKRYFLHTSAPLD 494  
 411 REETEVNTITTYDICTPRLKTOHNLTVSVNDNAPFESQTYTLRYRENNSPALH 470  
 495 YETREFRENVIVAVDSGSPSLSSKNSLTVKAGDTPNDPMEFGSVVEYYPENNIGERY 554  
 471 GSVATCRDSCANAQVYISLPPHDPQLPLGSLVSTINADNGOLFALRSIDFELQAF 530  
 555 ATVATDADSGKNAELAYSL-----DSSV-MGTF-AIDPDGDLVTVVTDREOTDRE 608  
 531 VGAADRSPPALSSQALRVLVADANDAP-----FVLYLQNGSAPCTELVRAEA 584  
 609 VNAKDKGIPVLOGSTVIYVADKNDNDPKFMQDVFIFYKEN-----LQPNSP-- 657  
 585 LVAKVAVVDGSGONAWLS 603  
 658 -VGMVTVMDADKGRNAEMS 675

RESULT 6  
 T00042  
 BH-protocadherin PCDH7 (clone BH-Pcdh-c) - human  
 C:Species: Homo sapiens (man)





Db 173 FRLLDAATDADVGENGYDQYEIVAGNVNKFRLVTTANPSGDTSYLHLETTGNLDRES 232  
 QY 205 VELLALITLGLDGGSPRPSGTLILILVDINDNPEVESIKKQVENSPLVTV 264  
 Db 233 KASYLINTISARGGSPRGLQVAVNTILDVNDNPIDHSDYNSINETALPGTPVTV 292  
 QY 265 SARLDGTGNGEIVAFATERTLKTFRIN-----STGSLHILKALINEAI---OTY 315  
 Db 293 MASDNDLDGNSKITV---YLAE-THQSTVNPETGIVISTEVENECQNTNKSASQSC 348  
 QY 316 TLTIAKDGCG--LSGKCTVVVHTVDINDNPELLM-----SSLTSPIPENSPE-TVAV 367  
 Db 349 VTTVARDHSGSPROGRTVTVNLDTDNDHPIISFRFPDQGVAVADENAVNCTVAA 408  
 QY 368 FRIIRDSDGNNAKWCST--QDHPFVLKPSVENVYTLVTEALDKERTENITITVD 425  
 Db 409 VAVKSDSGSLNRTSVRIYSGNELGHFRLEADLHIVRNCVLDKEIKGNLTVVAMD 468  
 QY 426 LCPRLKTOHNLTVVSDVNDNAPTFESQTTVTLRVRENNSPALHIGVSATRDSCGANAQ 485  
 Db 469 QGTPARTTAHLIIVDYNVDNHEPFEKSEYSAVLSELAPGSEFVASITADEDTGVNAQ 528  
 QY 486 VVYSIL----- 491  
 Db 529 VHYDILSGNELKWFMSMDPLTGLITVTPLEIRIDVELSISARDGCPNPKFAYTQAKYI 588  
 QY 492 -----PP-----HDPLPLGSLV----- 504  
 Db 589 ILDENDEAPQFSQRONVTGLGEDAPPQITVALMTATDHD--QGTNCSVTFPALAPSVERTLP 647  
 QY 505 ---SINADNGOLFALRSLDFEALQAFEFVGAADRGSPA--LSQALVRLVAVADANAP- 559  
 Db 648 LQFALDALGQLTRPRPDREKMSQEIPIYINADGAPTPGATPVMINADVANDNPQ 707  
 QY 560 ----EVLVYLQNGSAPCTELVPRAAEGLYAKVAVADSGSQANMLSTQILKATEPGLE 615  
 Db 708 FYPRIYVSLTDDDDIK--LKKEVEKERILHLVTAASDKDDGNALIEYRLESIGE-GLF 764  
 QY 616 GWANHGEV 624  
 Db 765 QLDARSGAI 773

## RESULT 8

T1419  
 seven-pass transmembrane receptor protein precursor - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C:Accession: T1419  
 R:Hadjantonakis, A.K.; Formstone, C.J.; Little, P.F.R.  
 submitted to the EMBL Data Library, October 1997  
 A:Description: The Celser family of novel evolutionarily conserved seven-pass transmembrane  
 A:Reference number: Z17881  
 A:Accession: T1419  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-3034 <HMD>  
 A:Cross-references: EMBL:AF031572; NID:g3800735; PID:g3800736; PIDN:AAC68836.1  
 C:Genetics:  
 A:Gene: Celser1  
 A:Map position: 15  
 C:Keywords: transmembrane protein  
 F:1-26/Domain: signal sequence #status predicted <STC>  
 F:27-3034/Product: seven-pass transmembrane receptor protein #status predicted <MNT>

## Query Match

Best Local Similarity 18.2%; Score 584; DB 2; Length 3034;  
 Matches 180; Conservative 95; Mismatches 245; Indels 80; Gaps 17;

QY 66 IVSDERTIGFLNPLTGLDILNKLKDLREEL-----CGTETECVLPFQULLLEKPRQ 115  
 Db 513 IVSGNLKGFYHLSGSDLVNPLDPEAIRYTLRIKADGGRPPLINSSGLV----- 566

QY 116 IFRAELVNRDINDHSPVELDREITLNLTESTPGATFLESADSDVGINNLNNTYI--- 172  
 Db 567 ----SVQVLDVNDNAPLFVSPPQAVLEVPVPLGHVSLHILQANADAGARARQYLVPI 622  
 QY 173 -----SSNVFHLINVD--NCGNVSELYLDKVLDRNEPELRLTLTGID 216  
 Db 623 ASTIVGSSVDSSENPASAPDFEPOIHNSGMITVCAE-----LDREVEHVSFGEAVD 676  
 QY 217 CGSPRPGTLLIRILVLDINDNVEFEVESLYKQVNPNSPGLVTVASRDLDTGSNCE 276  
 Db 677 HGSPAMSSASVSTLVLDVNDNDMPFTQPYELRLNEDAAVGSVLETLRARDNANS--- 733  
 QY 277 IVYAFATERTLKTFRINSTGN--LHLKALINEYAIQTYTLTIAKDGGLSGKCTVV 334  
 Db 734 -VITYQLTGCTRRFRALSSQSGGLTLLALPDYKQERYVLAVASD--GTSHSHAQVE 791  
 QY 335 VHYTDINDNPELLMSSLTSPIPENSF-ETVAVAFRIIRDSDGNNAKWCSTODHL-PV 382  
 Db 792 INVTDAMTHRPVFGSSHVYTSVSEDRPVGTISATISATDEDTENARITVLEDPVPOFR 851  
 QY 393 LKPSVENFYLTVMERALDREREYNTITVTDLGPRLKTQHNLTVYSDVNDNAPTS 452  
 Db 852 IDPDTGITYM-TE--LDYEDQAAVYTLTAQDNGIPQKSDTTSLELILIDANDNAPRL 908  
 QY 453 QTTTLRVRENNSPALHIGVSATRDSCGANAQVYSLRPPHDPOLPLGSLVSINADNG 512  
 Db 909 RDEFGVSFEEDAPRSTSVLQVSATRDSCGNRLRYTFQGGDGD--GDFF-1EPTSGV 964  
 QY 513 LPARSLDFEALQAFEFVGAADRGSP-AISSQALVRLVAVADANAP-----EVLVYL 565  
 Db 965 IRTORRLDRENVAVYINMALAVDRGSPNPLSASGIVSVLIDINDNPEVEKDELEVE 1024  
 QY 566 QNGSAPCTELVPRAAEGLYAKVAVADSGSQANMLSTQILKATEPGLEFVWANHGEV 625  
 Db 1025 EN-----SPGVSVAIRANDPDEGPAQIITVQIVEGVEYVQDILLSDLR 1072

## RESULT 9

T20968  
 hypothetical protein F15B9.7 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T20968; T26278  
 R:Percy, C.  
 submitted to the EMBL data library, August 1996  
 A:Reference number: Z19351  
 A:Accession: T20968  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-2610 <MIL>  
 A:Cross-references: EMBL:Z78013; PIDN:CAB01427.1; GSPDB:GN00023; CESP:F15B9.7  
 A:Experimental source: clone F15B9  
 R:Baynes, C.  
 submitted to the EMBL Data Library, August 1996  
 A:Reference number: Z20187  
 A:Accession: T26278  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-2610 <M12>  
 A:Cross-references: EMBL:Z78018; PIDN:CAB01449.1; GSPDB:GN00023; CESP:F15B9.7  
 A:Experimental source: clone W07G4  
 A:Gene: CESP:F15B9.7  
 A:Map position: 5  
 A:Introns: 169/2; 217/3; 239/3; 284/3; 515/2; 550/3; 738/3; 810/1; 838/3; 890/2; 977/

## Query Match

Best Local Similarity 16.9%; Score 543.5; DB 2; Length 2610;  
 Matches 172; Conservative 100; Mismatches 271; Indels 115; Gaps 17;

QY 36 AEETRGFTLANLADLGLV-----EELSARG---CIVSDERTIGFLNPLTGLDILLN 87  
 Db 376 ASDSYNVTILENTIPAVIAIVAKTDEDEGTNGKVVHSMASSSGIGLITDYSTGEVTLR 435

QY 88 EKLD-----REELCGPTPCVLPFOLLLEKPFQIFRAELIWDINDHSVPFLDR 136  
 Db 436 ERIIDAKNSPIYAVIRAKDAQPALSSVPLTIN-----VIDINDHAPITLIAA 482  
 QY 137 ETLTLESTTPGATFLLSADSVGGINNLRNTISSNVFHNHNDGNGANNVSELVL 196  
 Db 483 QAMITLEENVAIGEVEGVAVYIDEDSGPNITIKSMESSEDFIID-EDSG-----LTKT 535  
 QY 197 DKVIDREVEPELRLTLTGIDGSPRSGTTLIRLIVDINDNVEPEVESIKYQVPENSP 256  
 Db 536 TKLIDRETTARYSLKVTARDKGTPLNNTIYAVLADINDNAPTEFKKENTYISEMP 595  
 QY 257 VGSLLVTVSADDLTGSGEIVYAFYATERTLKTFRINSTGNLHLKALNEPAIOTYT 316  
 Db 596 RGSQITLTKA---VDNDEOKITRYRIEADREVFSILDIGDGAITSVSGELKRQD-HKVR 652  
 QY 317 LTIOAKDGGISGCTVYVHTDI-----NDNPELLMSLTSPIDENSP-ETVV--- 365  
 Db 653 VEISTDGGIAGKCVAVNFIDVNSAPYFNDHP-----FSVKIPEHSPIGYPIVITLK 705  
 QY 366 -----AVFIRIDRDSGNNAKMCISIODHLPEVLKPS 396  
 Db 706 VSFSEYFEGYGLKHFYVSSNLKKRIPNSANWMAEDHRODNNAIYVSISSQEFFRIDPS 765  
 QY 397 VENFTYLTTERALDRERTENITITVDTGLPRLKTQHNLVTVSVDNNAPEFSQTTY 456  
 Db 766 SGD---TSVSSDLDRERATFVITVTDHASPPLNTSTOIEVLIDINDNSPQFTSSSY 822  
 QY 457 TLARRENNSPALHIGVSATDRDSCANAOYTSLLPDPQLPGSLVSTINADNGOLFAL 516  
 Db 823 AATISEDIPVGTSLQYSAIDADIGPNGIYDYL-NESSSSISQILFRIDRTSGTLRVS 880  
 QY 517 RSLDFEALQAFEEFYGAADRGSPALSSQALVRYLVADANDAP-----FVLYPLONGSA 570  
 Db 881 SKLDREFAVYVLPFIPARDGTPLSLASSETITLILSDVNNNAPEFQSLDYIAEN--- 937  
 QY 571 PCTELVPRAEAGYLVAVAVDGSQONAMLSYOLLAKEPGLFCV---MAINGEVR 625  
 Db 938 -----SPVGSTVGTIVARADBDGDNADISFRIFGADAKLFIEEDAEQNGYVR 986

## RESULT 10

T00252  
 MEGF1 protein - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 21-Jul-2000  
 C:Accession: T00252  
 R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.  
 Genomics 51, 27-34, 1998  
 A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs  
 A:Reference number: Z14126; PMID:98360089; PMID:9693030  
 A:Accession: T00252  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-431 <NAK>  
 A:Cross-references: EMBL:AB011527; NID:93449285; PIDN:BA432458.1; PID:93449286  
 A:Experimental source: brain  
 C:Genetics:  
 A:Gene: MEGF1  
 C:Superfamily: rat MEGF1 protein; cadherin repeat homology; EGF homology; laminin G repeat  
 F:3953-3985/Domain: EGF homology <EGF>  
 F:3992-4023/Domain: EGF homology <EGF>

## Query Match

Best Local Similarity 16.9%; Score 542; DB 2; Length 4351;  
 Pred. No. 6.8e-27;  
 Matches 178; Conservative 99; Mismatches 238; Indels 160; Gaps 20;

QY 65 RIYSDETIGFLLNPLTGLLNEKIDRELGPTPCVLP-FOLLLEKPFQIF----- 117  
 Db 1803 KILPEALKFFKIDPSMGTLTTSTELDFED-----TFLFQ-----FNIYVHDGQT 1847

QY 118 -----RAELWVDINDHSPVELDREITNILESTTPGATFLLSADSVGGINNR 168  
 Db 1848 PILFARSAKAVITHAVDNDSPPRESEQIYEAVAEPIHPHGMLTLVQAEQDSDRV---- 1903  
 QY 169 NYTI-SSNVFHNVA-DNDEGNVS-----ELVLDKVD 201  
 Db 1904 TYSIKTSNDEAVTIPHTGQISVNPATRLRFQKFSIRASDGLYHDPAVVISLTVLD 1963  
 QY 202 R-----EEPELRLT-----LTGLD----- 216  
 Db 1964 KSLQFDQDVARAVENTENTPRKALVILGVGHNLNDLTSLFLLNGTDLPHMIESAGVLQT 2023  
 QY 217 -GGS-----PPRSGTTLIRLIVDINDNVEPEVESIKYQVPENS 255  
 Db 2024 RGSFTPREODDHEVAEVRDRNRVQRVAQALYRVSEVDNNTPEFOLPYTVIYDGT 2083  
 QY 256 PVGSLVTVSADDLTGSGEIVYAF--YATERTLKTFRINSTGNLHLKALNEPAIQ 313  
 Db 2084 EPGDVLFQVSARDKDKDGANGSVYGFADYA-----YFRIDPYVGDISLKKPFQYQALN 2137  
 QY 314 TYTLTQAKDGGISGCTVYVHTDINDNPELLMSLTSPIDENSPETVAVAFRIIDR 373  
 Db 2138 KHLKRYIANDSGIPPLQTEVEVHTVRKNSNPLFQSPYKAVKVPNT--TLTPTLHTQA 2195  
 QY 374 DSGNNAKMCISIODHLPEVLKPSVENEFTLVTERALDRERTENITITVD--LGTPL 431  
 Db 2196 RSPEGLRLIYINVEEPLMLFTTDFKTVGLVYTGDLDESKKHVYRAIDTALGS--- 2252  
 QY 432 KQHNLTVTVSVNNAPEFSQTTYTLKVRNNSPALHIGVSATDRDSCANAOYTSLL 491  
 Db 2253 FSEATVEVLEDINDNPPFESQVLTYSSESQATQVFIOLLASDQSGONOVSYQIV 2312  
 QY 492 PPHDPOLPLGSLVS-----INADNGOLFALRSIDFALAFERFVGAARSPALSSQALY 547  
 Db 2313 ED-----GDSVSKAFPRNGSTGEITFTQELDYTHGHFKYKRAADKGPPLTGETLV 2365  
 QY 548 RVLVADANDNAPFVLYPLONGSAPCTELVPRAEAGYLVAVAVDGSQONAMLSYOL 607  
 Db 2366 VVNVSDINDNPPKPREPOYAN-----VSELATCGHLVLKQALDPIGTSRLEVIL 2419  
 QY 608 KATEPGLFGVMAHG 622  
 Db 2420 SGNDRHFSINSTSG 2434

## RESULT 11

T15276  
 hypothetical protein R10F2.1 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C:Accession: T15276  
 R:Du, Z.; Gattung, S.  
 submitted to the EMBL data library, May 1997  
 A:Description: The sequence of C. elegans cosmid R10F2.  
 A:Reference number: Z18320  
 A:Accession: T15276  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-2163 <DUZ>  
 A:Cross-references: EMBL:AF003388; NID:92088850; PID:92088852; PIDN:AB454266.1; GSPDB  
 A:Experimental source: strain Bristol N2; clone R10F2  
 C:Genetics:  
 A:Gene: CESP.R10F2.1  
 A:Map position: 3  
 A:Insertions: 34/1; 114/3; 735/3; 1552/2; 1800/3; 1865/3; 1911/3; 2026/3; 2108/2; 2150/3

## Query Match

Best Local Similarity 15.1%; Score 485; DB 2; Length 2163;  
 Pred. No. 1.4e-23;  
 Matches 149; Conservative 82; Mismatches 224; Indels 60; Gaps 16;

QY 123 VRDINDSPVFLDREITNILESTTPGATFLLSADSVGGINNLRNTISSNVFHN 181  
 Db 46 LKDYDNMWPFIYFNEYHLYVREGPKPEEPILVAVASDMSDGTGEVSYHLTSSSSFSIN 105

QY 182 VHDNENGVSELYDKYLDEEVEBELRLTLGLDGGSPRSGLTLRLILVDINDNPE 241  
 Db 106 -----PVYGEVFAOKPLAR---GRFLHYVSAKKGQSSNSNHNITVYIDDKTRPQ 155  
 QY 242 EVESLYKVOVPENSPVSGSLVTVSARDLDTGSGNGEIVAFYAFVETRLKTERINSTGNL 301  
 Db 156 FERSKYETRTEDILPGIAIGSV-----GAMSKVKYSIXSGDD--HNFSIDEDTKI 206  
 QY 302 HUKALENTAIOITYTLTQAKDGGLSGKCTVYVAVHTVINDNPE-----LMSLITSLIP 357  
 Db 207 YVTRYLDAVDHDTVLNTQATLPSESNQOTVYFIEHDNDNPPQSSLSLEISREDLK 266  
 QY 358 ENSPTEVAVFRIRDRDSGNNAKVCISIODHLFVLKPSVE-NEYT--LVTERALDRER 414  
 Db 267 LHEPFYV---QASDKDKKNGEVKYSIISHP---GSSIELDOTGGLSGSLDYAT 320  
 QY 415 TEYNTITVTDLGPRLKTOHNLVTVSDVDNDAFPESQTYTLTVRENSPALHIGSVS 474  
 Db 321 RNYKLRVKATDGLIPRSMNTLFTHVLDVNDNNADEFEKSWTMEVLNSPKTIVGRKV 380  
 QY 475 ANDRDSGAAQVYSLPLPHDPLGSLVSINADNGOLFALRLDEALQAEFFRGA 533  
 Db 381 ANDKOSLENGQIVY-----RITNGSEYFCIDKLGKLVTKSIDREVISHVDSIVA 432  
 QY 534 ADRGSPALSSQALVRLVADANDNAFVLYLPLONGSAPCTELVPRAEAG---YLVAKV 589  
 Db 433 EKGVPFRSSTASARISVLDVNDNP-----SGLSTPLVAVAGSPLSLAIGTI 481  
 QY 590 VAVDGDSCGNANLSTQOLKATEPGILFGVWANGEV 624  
 Db 482 VAMDPEKGLNGSVLYR--AQLOSLF-VKXSGDV 513

## RESULT 12

A47543  
 R-cadherin precursor - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 03-May-1994 #sequence\_revision 26-May-1994 #text\_change 22-Jun-1999  
 C:Accession: A47543; 155501  
 R:Hutton, J.C.; Christofori, G.; Chl, W.Y.; Edman, U.; Guest, P.C.; Hanahan, D.; Kelly, M.O. Endocrinol. 7, 1151-1160, 1993  
 A:Title: Molecular cloning of mouse pancreatic islet R-cadherin: differential expression  
 A:Reference number: A47543; MUID:940067164; PMID:8247017  
 A:Accession: A47543  
 A:Molecule type: mRNA  
 A:Residues: 1-913 <HUT>  
 A:Cross-references: EMBL:X69966; NID:9429111; PIDN:CAA9589.1; PID:9429112  
 R:Matsumami, H.; Miyatani, S.; Inoue, T.; Copeland, N.; Gilbert, D.; Jenkins, N.; Takei, J. Cell Sci. 106, 401-409, 1993  
 A:Title: Cell binding specificity of mouse R-cadherin and chromosomal mapping of the gene  
 A:Reference number: 155501; MUID:940095672; PMID:8270638  
 A:Accession: 155501  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-913 <RES>  
 A:Cross-references: GB:DI4888; NID:9457658; PIDN:BA03605.1; PID:9457659  
 C:Comment: Cadherins mediate calcium-dependent intercellular adhesion and are thought to C:Superfamily: cadherin; cadherin repeat homology  
 C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pro F:1-27/Domain: signal sequence #status predicted <SIG>  
 F:128-166/Domain: propeptide #status predicted <PRO>  
 F:167-913/Product: R-cadherin #status predicted <MAT>  
 F:167-721/Domain: extracellular #status predicted <EXT>  
 F:169-274/Domain: cadherin repeat homology <CR1>  
 F:244-249/Region: cadherin binding #status predicted  
 F:277-389/Domain: cadherin repeat homology <CR2>  
 F:300-304/Domain: calcium binding #status predicted <CR3>  
 F:392-504/Domain: cadherin repeat homology <CR4>  
 F:507-612/Domain: cadherin repeat homology <CR5>  
 F:613-721/Domain: cadherin repeat homology <CR5>  
 F:722-753/Domain: transmembrane #status predicted <TM>  
 F:754-913/Domain: intracellular #status predicted <INT>

F:870-885/Region: serine-rich  
 F:280,409,554,629,658,699/Binding site: carbohydrate (Asn) (covalent) #status predict

Query Match 14.1%; Score 453; DB 1; Length 913;  
 Best Local Similarity 27.6%; Pred. No. 5.1e-22;

Matches 150; Conservativity 73; Mismatch 200; Indels 120; Gaps 18;

QY 139 TLNLESTPGATFLESADSDVGI--NDRNTYSSNVEFHNVDNENGVSELYVD 197  
 Db 136 TVALDPEQPNNDTLTPMPQHOSSGLARKQKRWIIP-----INVENSNGPPQQLVLR 190  
 QY 198 KVLDRREEVPELRLLTLGLDGGSP-----RSGLTLIR----- 229  
 Db 191 RSDKNDIP--IRVSITGVGADOPMEYFNIDSMGKRYVTRPMDREBRASYHLRAVDM 249  
 QY 230 -----LIVLDINDNPEFVSLKQVYVPEVSGSLVYVVSARDLD--TGSNG 275  
 Db 250 NGKNVNPIDLTYIVIDNDRNPEFVNGVNGSVDESKPGTYVTVTADADSTANG 309  
 QY 276 EIVYAFYATER--TLKFRINSTGNL-HUKALENTAIOITYTLTQAKD-----GGGL 327  
 Db 310 MVRRIYIQPPQSPSQMMFTINSETGIVYVAGLDREKVOQYTVIQAATDMGNLNYGL 369  
 QY 328 SGKCTVYVHTDINDNPELMSLTSPFENSPETVYVAFRIHDRSGNNAMVCISID 387  
 Db 370 SNTATAIIVTVDNDNPEFTSTFAGEVPEENRLEIYVAMLYWDRD-----QP 418  
 QY 388 HLP-----FVLKPSVENFTVLTAEALDREERTENITVTDLG--- 427  
 Db 419 HSPWMNAVYRISGDPSCGHSVRTDPVTNCGMYVYKAAVDELRAPMLVYMSNQAFLA 478  
 QY 428 ---TPRLKTOHNLVTVSDVDNDAFPESQTYTLTVRENSPALHIGSVSATRDSGANA 484  
 Db 479 SGIMSFQSTAGVITVTDVNE--APYFSPNKKILRLREGVPAGATLTFEAVAPDRFMQ 537  
 QY 485 QVYTSLEPLPHDPLGSLVSNADNGOLFALRLDEAL-----QAEFFVGAADRSPA 540  
 Db 538 AVRSKTL--SDP---ANWLMHINTSNGOITTAALIDRESLYTKNNYEATFLAANDIPP 591  
 QY 541 LSSQALVRLVADANDNAFVLYLPLONGSAPCTELVPRAE---AGYLVAKVAVDGD 596  
 Db 592 ASGCTGQIYILIDINDNP-----QLPKKAOICERPGLMAINITIADADM 637  
 QY 597 GQN 599  
 Db 638 DPN 640

## RESULT 13

C38992  
 cadherin 4 precursor - human  
 N:Alternate names: R-cadherin  
 C:Species: Homo sapiens (man)  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 20-Aug-1999  
 C:Accession: C38992  
 R: Suzuki, S.; Sano, K.; Tanihara, H.  
 Cell Regul. 2, 261-270, 1991  
 A:Title: Diversity of the cadherin family: evidence for eight new cadherins in neuron  
 A:Reference number: S24305; MUID:91283540; PMID:2059658  
 A:Accession: C38992  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-916 <SUZ>  
 A:Cross-references: GB:I34059; NID:9506409; PIDN:AAA35627.1; PID:9506410  
 C:Genetics:  
 A:Gene: GDB:CDH4  
 A:Cross-references: GDB:622850  
 A:Map position: 16q24.1-16qter  
 C:Superfamily: cadherin; cadherin repeat homology  
 C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane F:1-27/Domain: signal sequence #status predicted <SIG>  
 F:128-166/Domain: propeptide #status predicted <PRO>  
 F:167-913/Product: R-cadherin #status predicted <MAT>

F:167-721/Domain: extracellular #status predicted <EXT>  
 F:169-274/Domain: cadherin repeat homology <CR1>  
 F:244-249/Region: cadherin binding #status predicted  
 F:277-389/Domain: cadherin repeat homology <CR2>  
 F:300-304/Domain: calcium binding #status predicted <CAB>  
 F:392-504/Domain: cadherin repeat homology <CR3>  
 F:507-612/Domain: cadherin repeat homology <CR3>  
 F:613-721/Domain: cadherin repeat homology <CR5>  
 F:722-753/Domain: transmembrane #status predicted <TM>  
 F:754-913/Domain: intracellular #status predicted <INT>  
 F:870-885/Region: serine-rich  
 F:283,412,557,632,661,702/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.9% Score 445.5; DB 2; Length 916;  
 Best Local Similarity 26.2%; Pred. No. 1.6e-21;  
 Matches 157; Conservative 81; Mismatches 210; Indels 151; Gaps 21;

QY 129 HSPVFLDREITLNLLESTPGATFLLSADSDVGI-NLNNTYISSNVYFHINVDNGE 187  
 DB 131 HKPQKGVVLD--PSPPKDTLLPMPQHONANGLRRKRKDWYTP-----INVPENSR 183  
 QY 168 GNVYSELVLDKYLDEEPELRLTLTGIDGSSP-----RSGTTLIR----- 229  
 DB 194 GPFPQQLVIRSDKNDIP-IRYSITGVGADQPMVEVFSINSMGRMYVTRPMDREHAS 242  
 QY 230 -----ILVLDINDVPEFVESLYVQVPEVSPVGLVVTYSAR 267  
 DB 243 YHLRAHVDNMGKNVKNPIDLYIVDMNDNHPFINOVNCSDESKPTVMTITAN 302  
 QY 268 DLD--TGSNGEIVYAFYATER--TLKTRINSTGNL-HLKAELENIQIOTYTLTIOAK 322  
 DB 303 DADSTTANGVRYRIYQTPQSPQNMFTINSEIGDIYVAAQMDREKVOQYIVYOAT 362  
 QY 323 D-----GGSLGSKCTVYVHTVDINDNPELLMSLTSPINSEVTVVAVFRIDRSGN 377  
 DB 363 DMENGLNTGSLNTAIIITVDVNDNPESTAFAGEVNSVETVAVMLTVMDRD--- 419  
 QY 378 NAKWVCSIODHLP-----FLKPSVENFYTLTEKALDREERTENIT 420  
 DB 420 -----QHPSPMMNAVYRIISDPSGHSVRDPTNMGMYVYKAAVDYELNRAPMLT 471  
 QY 421 ITYIDLG-----TPRLKQHNLTFTVSDVNDNAPTESQTYTLRVRENSPALHIGSVS 474  
 DB 472 VMSVNAQAPLASGIOMSFQSTAGVTISIMDINE-APYFSPNHKLRLEEGVPPTVLTFS 530  
 QY 475 AFDDSGANAOVTSLRPHDPQLPLGSLVSIANNGOLFALSLDPEAL-----QAEFR 530  
 DB 531 AVDPDRMQQAVRYSKL--SDP-----ASWLHINATNGOITTVAVLDRESLYTKNNYEAT 584  
 QY 531 VGAADRGSPALSSQALVTVLVADANDNAPFVLYPLONGSAPCTELVRAAE---AGYLV 586  
 DB 585 FLADNGIIPRPSGCTGLQIYLIDINDNAP-----ELLPKEAQICERPLNA 630  
 QY 587 AKVAVVDCD-----SGQNMALISTQLKATEPGLFEGV 617  
 DB 631 INTAADADVHPNIGPYFELPVPVPAVRKNMTITRLNGDYAOLSLKIL-YLEAGMVDV 688

## RESULT 14

S44887  
 ZK112.7 protein - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 14-Sep-1994 #sequence\_revision 12-May-1995 #text\_change 23-Mar-2001  
 C:Accession: S44887  
 R:Du, Z.  
 submitted to the EMBL Data Library, May 1993  
 A:Description: Sequence of the C. elegans cosmid ZK112.  
 A:Reference number: S44616  
 A:Accession: S44887  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-3343 <DU>  
 A:Cross-references: EMBL:114324; NID:g289740; PID:g289742

C:Genetics:  
 A:Intons: 39/1; 74/3; 122/3; 209/2; 287/3; 333/2; 489/1; 611/2; 714/1; 939/3; 1002/2  
 C:Keywords: cytoskeleton; transmembrane protein

Query Match 13.8% Score 442.5; DB 2; Length 3343;  
 Best Local Similarity 26.5%; Pred. No. 1.8e-20;  
 Matches 175; Conservative 94; Mismatches 249; Indels 143; Gaps 28;

QY 35 VAESTERTFLANLAIDLGLGVEELSARGCIYSDETIGLLNPLTGLDLLNEXKIDRE 94  
 DB 1564 ISEIADFTTEVGIYI-LGAGLEGSVF-----IQDDYNF-TISPDDGIFITNSPLDEFEN 1617  
 QY 95 LCGTEPCVLPFQLLLEKPEQIFRAELWVDINDSPVELDRE-TITNILES-TTPGATF 152  
 DB 1618 -----IKTYRFTNACKSTQVLIHTVDEDEAPRFITGVVNLKYLEELDVPYPL 1669  
 QY 153 LLES--AHSDVGINNLRNTTSS--NVFHNVDHNGGNNYSELVLDKYLDEEVPPL 208  
 DB 1670 ITGSSIAEDDEGQNLVYTSILSGNTSLFAVN--STTGDIISL-----PLDRESSLA 1722  
 QY 209 RLNLITGIDGSPSPSGTTLIRILVLDINDVPEFVESLYKQVPEVSPVGLVTVYSAR 268  
 DB 1723 ELIEAKDAGIPSLATSKILIHVGIDINDNTPPELISFYIKISENSKIGSKILILATD 1782  
 QY 269 LDTGSGEIVYAFYATERLTFTFRINSTGNLHLKAELENIQIOTYTLTIOAKDGGIS 328  
 DB 1783 KD--KDAELOYSLESNDEITLI--PFRINAVATGMITVAGKNRENEBFRFVAVTD-GEKS 1838  
 QY 329 GKCTVYVHTVDINDNPEL--LMSSTSPINSEPTVAVFRIDR-----SGNNA 379  
 DB 1839 SKYIVIEIHVEDFNHPIINDRNSDLEVPDPTRYVE-IIHIVNVDLSDHLKFLSNLS 1897  
 QY 380 KWCYS-----IOD-----HLPF----- 391  
 DB 1898 NNLNLSNGEITLKSPLQTAVPVRYVSDDAGVAFMEYLFHPSKHPVPEKIDTVSY 1957  
 QY 392 -----VLRPSVENF-YTLV-----TERALDREERTENI-TTV 423  
 DB 1958 REHDEQELAFVRANGSISYSIVSRCSHLENEKSTGLIKTKSSIDAEYSCELVFIAT 2017  
 QY 424 TDLGTPRLKQHNLTFTVSDVNDNAPTESQTYTLRVRENSPALHIGSVATDSCAN 483  
 DB 2018 TYEDKNPLETITKATIKIYINDNSPRDQQLRYRNVENSQPKL-IGHYLA--RDIDRS 2074  
 QY 484 AOYVTSIL--PPHDQPLGSLVSIANDNGOLFALSLDPEALQAEFFRYGAADRGSPA 540  
 DB 2075 SRVFEIIVGDANHEFV-----TESQIESVRDLDRKSEYHLIYEALIDGKPR 2125  
 QY 541 LSSQALVRYLVADANDNAPFV--LYPLONGSAPCTELVPRAEAGYLVAKYVAVDGDGSG 598  
 DB 2126 RRGNTTVIVTVLDEDDNAPRSRIFHE-----VPEVRIIGEPIVQLSADSDAHERS 2176  
 QY 599 N 599  
 DB 2177 N 2177

## RESULT 15

G02678  
 cadherin-14 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 21-Jan-2000  
 C:Accession: G02678  
 R:Shibata, T.; Shimoyama, Y.; Gotoh, M.; Hirohashi, S.  
 submitted to the EMBL Data Library, May 1996  
 A:Reference number: H01584  
 A:Accession: G02678  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-790 <SHI>  
 A:Cross-references: EMBL:U059325; NID:g1389852; PIDN:AA02933.1; PID:g1389853  
 C:Superfamily: cadherin; cadherin repeat homology  
 F:163-268/Domain: cadherin repeat homology <CDH>

Query Match	13.7%;	Score 438;	DB 2;	Length 790;
Best Local Similarity	28.4%;	Pred. NO. 4e-21;		
Matches 154;	Conservative 84;	Mismatches 204;	Indels 100;	Gaps 21.

QY	66	IVS0ETIG-FLLPMLGDLINKLDRRELIC-----GPTPCVLPQOLL	110
Db	91	ILTBEGACTITIIDDTGGDTHSTKSLDREOKTHVLAQAIDRTRKFLDP-----	141
QY	111	EKPQIFRAELMWADINDHSPFLDREITINILESTTPGATFLESAMD--VGINL	167
Db	142	ESEFII-----KVQDINDNAPKFTDGPYIVTPEMSDMGTSLQVATDADDPYGSAR	196
QY	168	RNYTI-SSNYFHLNVADNGBGNYSLSVLVDKVLDRFEVELRLITLGLD----GCSP	221
Db	197	VVYSILLOGPYFSY--DPKTGVIRTLA---HNMDRAREHYSVIOAKMAGOVG--	247
QY	222	RSGTTLRLILVDINDVPEFVESLYVOYQENSPVGLVVTYSARDLDGSGEITYA-	280
Db	248	LSGSTYVITLTDVNDNPPRPQAHYLDYLPESQVSGSAVGKIKANDADGSNADMTYI	307
QY	281	-----FVYATERTLTKFRINSTGNLHKALNEAIOYTLTIOAKD-----	324
Db	308	INGDGMGFISTDK-----EIRREILSLKRLKLANEKKKSTLNEGANTHLDPRFSH	360
QY	325	-GGISGKCTVYVAVHTDINDNPELMLSSLTSPIDENP-ETVAVAFIRDRDGSNNAKV	382
Db	361	LGPBKDATMLKIIYGDV-DEPPLFSMPSYLMEEVENAKICTVVGTVLAOPDSTNSL--	416
QY	383	CSIDHLEPVLAKPSVE-----NNTYLTVEALDRERTENTINITYTDLGTPRKT	433
Db	417	-----VRYFINVYEDDREPFNIDANGTIRTKTVLVDRETPWNTIVTASEINDPDL	470
QY	434	QHNLTVYSDVDNDAPEFSQTYTYLRYRENNSPALHIGSATDRDGSAAQVYTSLLP	493
Db	471	HVYIGIRLVDNDNPPELAR-EYLIICVENSMPQVYHTISATDKDFANGPRNFEL--	527
QY	494	HDPOLPGSLVST--MADNGQLFRLSLDPE--ALQAFERRVGAADRGSALLSQALRV	549
Db	528	-DERLPVAPNFTLKDNEEDNASILTRRRRRSRQVDYVYLPIMISDGIPTLSSSTLTI	586
QY	550	LY 551	
Db	587	RV 588	

```
Search completed: May 16, 2003, 10:47:42
Job time : 30 secs
```



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 16, 2003, 10:46:02 ; Search time 37 seconds  
(without alignments)  
3491.663 Million cell updates/sec

Title: US-10-050-704-125

Perfect score: 3208

Sequence: 1 MEARVHALOKRQVSLCVF.....KATEPGLFGWAHNGEVRTA 627

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.rodent:\*  
12: sp.virus:\*  
13: sp.vertibrate:\*  
14: sp.unclassified:\*  
15: sp.rvirts:\*  
16: sp.bacteriap:\*  
17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2536	78.7	723	11	Q925L6
2	2536	78.7	786	11	Q91Y04
3	2256	70.3	799	11	Q91X29
4	2234	70.3	800	4	Q96T99
5	2230	70.1	844	11	Q925L1
6	2240	69.8	776	4	Q96SE9
7	2167	67.5	799	11	Q91Y00
8	2164	67.5	784	11	Q91X27
9	2159	67.3	794	11	Q91X28
10	2157	67.2	784	11	Q925M6
11	2157	67.2	794	11	Q925L0
12	2155	67.2	792	11	Q91Y02
13	2148	66.6	796	11	Q91Y05
14	2138	66.6	792	11	Q925L3
15	2104	65.6	799	11	Q91YD8
16	2098	65.4	799	11	Q925L4

17	2078	64.8	792	11	Q91X25
18	2052.5	64.0	734	4	Q96T90
19	2050.5	63.9	774	11	Q91Y48
20	2042.5	63.7	798	11	Q925M2
21	2040.5	63.6	793	11	Q91X21
22	2036.5	63.5	798	11	Q91X23
23	2023.5	63.1	772	11	Q91X24
24	2023.5	63.1	796	11	Q91Y06
25	2013.5	62.8	779	11	Q91X22
26	2011.5	62.7	797	11	Q91Y28
27	2009.5	62.6	789	11	Q91Y07
28	2007	62.6	801	11	Q925L2
29	2007	62.6	801	11	Q91Y01
30	1996.5	62.2	797	11	Q91X26
31	1986.5	61.9	794	11	Q91Y05
32	1977.5	61.6	802	11	Q91Y03
33	1969.5	61.4	784	11	Q925M5
34	1879	58.6	744	4	Q9HAB4
35	1836	57.2	587	4	Q96T98
36	1723.5	53.7	818	11	Q91Y08
37	1693	52.8	593	11	Q925M0
38	1573	49.0	595	11	Q925M7
39	1432.5	44.7	935	11	Q91XY1
40	1428.5	44.5	931	11	Q91XY6
41	1423	44.4	932	11	Q91X29
42	1418	44.2	813	4	Q9Y5D2
43	1418	44.2	931	4	Q9Y5G8
44	1415.5	44.1	850	4	Q9Y5E0
45	1415.5	44.1	936	4	Q9Y5H3

## ALIGNMENTS

## RESULT 1

ID Q925L6 PRELIMINARY: PRT: 723 AA.  
AC Q925L6  
DT 01-DEC-2001 (TREMBLER, 19, Created)  
DT 01-DEC-2001 (TREMBLER, 19, Last sequence update)  
DT 01-JUN-2002 (TREMBLER, 21, Last annotation update)  
DE Protocadherin-beta (Fragment).  
GN PCDB15 OR PCDB17.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21223055; PubMed=11322959;  
RA Vanhaest K., Kools P., Vanden Eynde E., van Roy F.;  
RT "The human and murine protocadherin-beta one-exon gene families show  
RT high evolutionary conservation, despite the difference in gene  
RT number.";  
RL FEBS Lett. 495:120-125(2001).  
CC -i- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.  
DR EMBL: AF326308; AAK53233.1;  
DR MGD: MGI:2136750; Pcdh15.  
DR InterPro: IPR002126; Cadherin.  
DR Pfam: PF00028; cadherin: 5.  
DR PROSITE: PS00232; CADHERIN\_1; UNKNOWN\_5.  
DR PROSITE: PS0268; CADHERIN\_2; 6.  
KW Calcium-binding; Cell adhesion; Glycoprotein.  
FT NON\_TER 723  
FT SEQUENCE 723 AA; 79555 MW; 678543AB3647DEB6 CRC64;

Query Match 78.7%; Score 2526; DB 11; Length 723;  
Best Local Similarity 78.1%; Pred. No. 8.5e-167;  
Matches 489; Conservative 57; Mismatches 80; Indels 0; Gaps 0;  
OY 1 MEARVHALOKRQVSLCVF...KATEPGLFGWAHNGEVRTA 627  
DB 1 MEARVHALOKRQVSLCVF...KATEPGLFGWAHNGEVRTA 627

```

QY 61 ARGCRTVSEDTIGFLINPLTGDLLNEKLDREELCGTEPCVLPFOILLEKPFQIFRAE 120
DB 61 AREARIVSQNTRFLLSPILSTGELLINKEKLDREELCGTEPCVLPFOILLERPPQIYRAA 120
QY 121 LWRVINDHSPVFLDEITLNLTESTPGATFLESADSDVGINLKNYITSSVNFYHI 180
DB 121 LHKINDNSPVFLDEIKELIKISESTPGATLLERADADVGNLSNYSITSPNDYFHI 180
QY 181 NVHDNGEGNVSELYDKVLDREVEPELRLTLTGIDGSPRSSTGLTRILVLDINDVP 240
DB 181 NVHDNGEGNIPRLTQHTITVOVSDINDNAPFTQTSYTMFEVRENNSPALHIGTISATSDS 240
QY 241 EVELSELYKQVSPENPVGSLVTVASARDLDGSGEIVYAFYATERTLKTFRINSTGN 300
DB 241 QFVQSLKYQSPENPVGSLVTVASARDLDGSGEIVYAFYATERTLKTFRINSTSGE 300
QY 301 LHKAELENEAIQIYTLTIQAKDGGLSGKCTVVVHVTDINDNPPELLMSLTSPIPENS 360
DB 301 LYLKRELNEAIQIYTLTIQAKDGGLSGKCAVVEVADVNDNPEFLLSLNSPIPENS 360
QY 361 PETVAVAFRIRDRSGNNAKKWCSTQDHLPEVLKPSVENFTYTERALDREERTENIT 420
DB 361 QETVAVAFKIRDRSGNNKGTICSTANDLPVLKPSVENFTYTERKPLDRESNTEYNT 420
QY 421 ITVDTLGPRLKTOHNLTVVSDVNDNAPFTSQTYTTRVRENNSPALHIGSVATDRDS 480
DB 421 ITVDMGIPRLTQHTITVOVSDINDNAPFTQTSYTMFEVRENNSPALHIGTISATSDS 480
QY 481 GANAQVTSLSLPHDPOLPLGSLVSINADNGQLFALRSLDPEALQAEFFRYGADRGSPA 540
DB 481 GSNMHTYSLRPHDPOLALDLSISINADNGQLFALRLDPEALQAEFFRYGADRGSPA 540
QY 541 LSSQALVRLVADANDNAPFVLYPLONGSAPCTELVPRAAGYLVAKVAVDSDGSCNA 600
DB 541 LSSQALVRLVITDNDNAPFVLYPMONASAPYTELPRAAPGYLVAKVAVDSDGSCNA 600
QY 601 WLSYQLKATEPGLFVGVAHNGEVRT 626
DB 601 WLSFQLKATEPGLFVSVAHNGEVRT 626

```

RESULT 2  
Q91X04 PRELIMINARY; PRT; 786 AA.

AC Q91X04; PRT; 786 AA.  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Protocadherin beta 15.  
GN PCDBH15.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OC NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RX MEDLINE=99308636; PubMed=10380929;  
RA Wu Q., Maniatis T.;  
RT "A striking organization of a large family of human neural cadherin-  
like cell adhesion genes.";  
RL Cell 97:779-790(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RX MEDLINE=20202599; PubMed=10716726;  
RA Wu Q., Maniatis T.;  
RT "Large exons encoding multiple ectodomains are a characteristic  
feature of protocadherin genes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:3124-3129(2000).  
RN [3]  
RP SEQUENCE FROM N.A.

```

RC TISSUE-BRAIN;
RA MEDLINE=21154914; PubMed=11230163;
RX Wu Q., Zhang T., Cheng J.-F., Kim Y., Grimwood J., Schmutz J.,
RA Dickson M., Noonan J.P., Zhang M.Q., Myers R.M., Maniatis T.;
RT "Comparative dna sequence analysis of mouse and human protocadherin
RT gene clusters.";
RL Genome Res. 11:389-404(2001).
CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.
DR EMBL: AY013776; AKK26065.1; -.
DR MGI: 2136750; Pcdh15.
DR Interp: JPK002126; Cadherin.
DR Pfam: PF00028; cadherin; 5.
DR PROSITE: PS00232; CADHERIN_1; UNKNOWN_5.
DR PROSITE: PS50268; CADHERIN_2; 6.
KW Calcium-binding; Cell adhesion; Glycoprotein.
SQ SEQUENCE 786 AA; 86292 MW; 0EA6C04986F6FA90 CRC64;

Query Match 78.7%; Score 2526; DB 11; Length 786;
Best Local Similarity 78.1%; Pred. No. 9,6e-167;
Matches 489; Conservative 57; Mismatches 80; Indels 0; Gaps 0;

QY 1 MEARVYHALQKROVSLCYFLGVSAGAPLRYFAEETERTFLANLADIGLYEELS 60
DB 1 MEARLDQAVQKROVLFVYFLGASWASADQLQYSAEETERTFLANLAKDGLGELIS 60
QY 61 ARGCRTVSEDTIGFLINPLTGDLLNEKLDREELCGTEPCVLPFOILLEKPFQIFRAE 120
DB 61 AREARIVSQNTRFLLSPILSTGELLINKEKLDREELCGTEPCVLPFOILLERPPQIYRAA 120
QY 121 LWRVINDHSPVFLDEITLNLTESTPGATFLESADSDVGINLKNYITSSVNFYHI 180
DB 121 LHKINDNSPVFLDEIKELIKISESTPGATLLERADADVGNLSNYSITSPNDYFHI 180
QY 181 NVHDNGEGNVSELYDKVLDREVEPELRLTLTGIDGSPRSSTGLTRILVLDINDVP 240
DB 181 NVHDNGEGNIPRLTQHTITVOVSDINDNAPFTQTSYTMFEVRENNSPALHIGTISATSDS 240
QY 241 EVELSELYKQVSPENPVGSLVTVASARDLDGSGEIVYAFYATERTLKTFRINSTGN 300
DB 241 QFVQSLKYQSPENPVGSLVTVASARDLDGSGEIVYAFYATERTLKTFRINSTSGE 300
QY 301 LHKAELENEAIQIYTLTIQAKDGGLSGKCTVVVHVTDINDNPPELLMSLTSPIPENS 360
DB 301 LYLKRELNEAIQIYTLTIQAKDGGLSGKCAVVEVADVNDNPEFLLSLNSPIPENS 360
QY 361 PETVAVAFRIRDRSGNNAKKWCSTQDHLPEVLKPSVENFTYTERALDREERTENIT 420
DB 361 QETVAVAFKIRDRSGNNKGTICSTANDLPVLKPSVENFTYTERKPLDRESNTEYNT 420
QY 421 ITVDTLGPRLKTOHNLTVVSDVNDNAPFTSQTYTTRVRENNSPALHIGSVATDRDS 480
DB 421 ITVDMGIPRLTQHTITVOVSDINDNAPFTQTSYTMFEVRENNSPALHIGTISATSDS 480
QY 481 GANAQVTSLSLPHDPOLPLGSLVSINADNGQLFALRSLDPEALQAEFFRYGADRGSPA 540
DB 481 GSNMHTYSLRPHDPOLALDLSISINADNGQLFALRLDPEALQAEFFRYGADRGSPA 540
QY 541 LSSQALVRLVADANDNAPFVLYPLONGSAPCTELVPRAAGYLVAKVAVDSDGSCNA 600
DB 541 LSSQALVRLVITDNDNAPFVLYPMONASAPYTELPRAAPGYLVAKVAVDSDGSCNA 600
QY 601 WLSYQLKATEPGLFVGVAHNGEVRT 626
DB 601 WLSFQLKATEPGLFVSVAHNGEVRT 626

```

RESULT 3  
Q91X29 PRELIMINARY; PRT; 799 AA.

AC Q91X29; PRT; 799 AA.  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)



DE Protocadherin beta 20.  
 GN PCDB20.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RX MEDLINE=99308636; PubMed=10380929;  
 RA Wu Q., Maniatis T.;  
 RT "A striking organization of a large family of human neural cadherin-  
 RT like cell adhesion genes";  
 RL Cell 97:779-790(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RX MEDLINE=20202599; PubMed=10716726;  
 RA Wu Q., Maniatis T.;  
 RT "Large exons encoding multiple ectodomains are a characteristic  
 RT feature of protocadherin genes";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:3124-3129(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RX MEDLINE=21154914; PubMed=11230163;  
 RA Wu Q., Zhang T., Cheng J.-F., Kim Y., Grimwood J., Schmutz J.,  
 RA Dickson M., Noonan J.P., Zhang M.Q., Myers R.M., Maniatis T.;  
 RT "Comparative dna sequence analysis of mouse and human protocadherin  
 RT gene clusters";  
 RL Genome Res. 11:389-404(2001).  
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.  
 DR EMBL: AY013782; AAK26071.1;  
 DR MGD: MGI:2136758; Pcdh20.  
 DR InterPro: IPR002106; AAcRNA\_1ligase1.  
 DR Pfam: PF00028; cadherin. 5.  
 DR PROSITE: PS00179; AA\_TRNA\_LIGASE\_II.1; UNKNOWN\_1.  
 DR PROSITE: PS00232; CADHERIN\_1; UNKNOWN\_5.  
 DR PROSITE: PS50268; CADHERIN\_2; 5.  
 DR Calcium-binding; Cell adhesion; Glycoprotein.  
 SQ SEQUENCE 799 AA; 87800 MW; 99C7A595F1673147 CRC64;  
 Query Match 70.3%; Score 2256; DB 11; Length 799;  
 Best Local Similarity 68.9%; Pred. No. 5 3e-148;  
 Matches 426; Conservative 91; Mismatches 101; Indels 0; Gaps 0;  
 QY 9 LKROYSLCLVFLGYSWAGAEPLRYFAETRGTFIANLAIDLGLVELSARGCRIVS 68  
 DB 9 LKROYLVEFVFLGLSQASAEISLRYSAVABETELGSFVANLAKDLGLGVAELSSREARVVS 68  
 QY 69 DETIGLLNPLTGLDLLNEKLDRELCPTEPCVLPOLLLEKPPQIFRAELWVDIND 128  
 DB 69 DDKKHLNLNLTGLDMLNERLDRELCSGPCVLPFOVYENPQFRAELWVDIND 128  
 QY 129 HSNVFDREITLILSTPGATFELLSAHSQVGINNRYNTISSNVEFHINVDNGS 188  
 DB 129 HSTFELDKETITISSTIGATFLEKQDDIGSNLSQDYSISSTFFYKIHDSGCG 188  
 QY 129 HSTFELDKETITISSTIGATFLEKQDDIGSNLSQDYSISSTFFYKIHDSGCG 188  
 DB 129 HSTFELDKETITISSTIGATFLEKQDDIGSNLSQDYSISSTFFYKIHDSGCG 188  
 QY 189 NVYSELVLDKVDREVEPELRLTLGLDGSPPRSCTLLIRLVLDINDNVEFEVSLX 248  
 DB 189 KIYPELVLDKALDHEESELRLTLGLDGSPPRSCTLLIRLVLDINDNVEFEVSLX 248  
 QY 249 VQVPEPSVGLSVYVSARDLDTGSGNGEIVAFYATERTKTRFINSNGNLHKAELN 308  
 DB 249 VQVPEPSVGLSVYVSARDLDTGSGNGEIVAFYATERTKTRFINSNGNLHKAELN 308  
 QY 309 YEAIQYTLTIQAKDGGSGKCTGVVHTDINDNEPELLMSLTFPENSPELVAVF 368  
 DB 309 FEYIQSVTIQATDGGGSAKCTLSKYVLIDINDNPEYIMSVTATPEMASETLVALF 368  
 QY 369 RIRDRSGNNAKVCISIDHLFVLKPSVENFTYLTALDRERTEVNTITVVDLCT 428  
 DB 369 RIRDRSGNNAKVCISIDHLFVLKPSVENFTYLTALDRERTEVNTITVVDLCT 428

DB 369 SVRDSDGDNCRILCSIODLPLILKPSKFNFTLLSEKALDRERREAVNTITASDMCT 428  
 QY 429 PRKTOHNLIVTVSDVNDNAPEFSQTYTLRVENNSPALHIGSVATDRSGANDQVY 488  
 DB 429 PRLTQHTIRQVSDINDNAPAFQISYTFITENNSPALHIGSVATDRSGANDQVY 488  
 QY 489 SLPPHPDPLGLSIVSINDNGOLFALRSIDEALQAFERGVAGDRGSPALSSQALVR 548  
 DB 489 SLPPHPDPLGLSIVSINDNGOLFALRSIDEALQAFERGVAGDRGSPALSSQALVR 548  
 QY 549 VLVDANDNAPEVLYPQNGSAPCTELVPRAEAGYIVAVAVDQSGONANLSTOLK 608  
 DB 549 VLVDANDNAPEVLYPQNGSAPCTELVPRAEAGYIVAVAVDQSGONANLSTOLK 608  
 QY 609 ATEPGLFVNAHNGEVRT 626  
 DB 609 ATEPGLFVNAHNGEVRT 626  
 RESULT 4  
 ID Q96T99 PRELIMINARY; PRT; 800 AA.  
 AC Q96T99;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Protocadherin-beta10.  
 GN PCDB10.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=21223055; PubMed=11322959;  
 RA Vanhaelt K., Kools P., Vanden Eynde E., van Roy F.;  
 RT "The human and murine protocadherin-beta one-exon gene families show  
 RT high evolutionary conservation, despite the difference in gene  
 RT number";  
 RL FEBS Lett. 495:120-125(2001).  
 CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.  
 DR EMBL: AF217748; AAK51616.1;  
 DR InterPro: IPR002106; AAcRNA\_1ligase1.  
 DR Pfam: PF00028; cadherin. 5.  
 DR PROSITE: PS00179; AA\_TRNA\_LIGASE\_II.1; UNKNOWN\_1.  
 DR PROSITE: PS00232; CADHERIN\_1; UNKNOWN\_5.  
 DR PROSITE: PS50268; CADHERIN\_2; 6.  
 DR Calcium-binding; Cell adhesion; Glycoprotein.  
 SQ SEQUENCE 800 AA; 87537 MW; 7DF4B52E0B78A263 CRC64;  
 Query Match 70.3%; Score 2254; DB 4; Length 800;  
 Best Local Similarity 70.6%; Pred. No. 7 4e-148;  
 Matches 436; Conservative 70; Mismatches 112; Indels 0; Gaps 0;  
 QY 10 OKROYSLCLVFLGYSWAGAEPLRYFAETRGTFIANLAIDLGLVELSARGCRIVS 69  
 DB 10 OKROYLVEFVFLGLSQASAEISLRYSAVABETELGSFVANLAKDLGLGVAELSSREARVVS 69  
 QY 10 OKROYLVEFVFLGLSQASAEISLRYSAVABETELGSFVANLAKDLGLGVAELSSREARVVS 69  
 DB 10 OKROYLVEFVFLGLSQASAEISLRYSAVABETELGSFVANLAKDLGLGVAELSSREARVVS 69  
 QY 70 ETIGLLNPLTGLDLLNEKLDRELCPTEPCVLPOLLLEKPPQIFRAELWVDIND 129  
 DB 70 DDKKHLNLNLTGLDMLNERLDRELCSGPCVLPFOVYENPQFRAELWVDIND 129  
 QY 130 SPVFDREITLILSTPGATFELLSAHSQVGINNRYNTISSNVEFHINVDNGE 189  
 DB 130 SPVFDREITLILSTPGATFELLSAHSQVGINNRYNTISSNVEFHINVDNGE 189  
 QY 130 SPVFDREITLILSTPGATFELLSAHSQVGINNRYNTISSNVEFHINVDNGE 189  
 DB 130 SPVFDREITLILSTPGATFELLSAHSQVGINNRYNTISSNVEFHINVDNGE 189  
 QY 190 VYSELVLDKVDREVEPELRLTLGLDGSPPRSCTLLIRLVLDINDNVEFEVSLX 249  
 DB 190 VYSELVLDKVDREVEPELRLTLGLDGSPPRSCTLLIRLVLDINDNVEFEVSLX 249  
 QY 249 VQVPEPSVGLSVYVSARDLDTGSGNGEIVAFYATERTKTRFINSNGNLHKAELN 309  
 DB 249 VQVPEPSVGLSVYVSARDLDTGSGNGEIVAFYATERTKTRFINSNGNLHKAELN 309

Db 250 QAPENSPIGFLIVKWAEDVDGNAEVSYSFSDASENIRTFQINPESGEILFELLDY 309  
 QY 310 EAIQYTTLTQAKDGGGSGKCTVYVHTDINDNPELMSLSPSPENSPETVAVER 369  
 Db 310 ELVNSTKINIQAAMDGGGSGARCRVLEVLDTINDNPELIVSSFSNSAENSPEETPLAVEK 369  
 QY 370 IINDRDSGNNAKAVCSIOHLPVLKPSVENEFTLVTERALDREERTENITITVDLCTP 429  
 Db 370 IINDRDSGNNAKAVCSIOHLPVLKPSVENEFTLVTERALDREERTENITITVDLCTP 429  
 QY 430 RLKTHNLTIVSDVNDNAPFTFSQTTTLRVRENNSPALHIGSVSATDRSGANAOVYS 489  
 Db 430 RLKTHNLTIVSDVNDNAPFTFSQTTTLRVRENNSPALHIGSVSATDRSGANAOVYS 489  
 QY 490 ILPPDHPPLPLASIVSINADNGHLEFALRSIDYEALEAFEFVGCATDRSGSPALSSALVRY 549  
 Db 490 ILPPDHPPLPLASIVSINADNGHLEFALRSIDYEALEAFEFVGCATDRSGSPALSSALVRY 549  
 QY 550 LVADANDNAPFLVPLONGSAPCTELVPRAEAGYLVAAVVDGSGQNAWLSYQLKA 609  
 Db 550 LVADANDNAPFLVPLONGSAPCTELVPRAEAGYLVAAVVDGSGQNAWLSYQLKA 609  
 QY 610 TEPGLFGVWANGEVRTA 627  
 Db 610 TEPGLFGVWANGEVRTA 627

## RESULT 5

Q925L1 PRELIMINARY; PRT; 844 AA.  
 AC Q925L1;  
 DT 01-DEC-2001 (TREMblrel. 19, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Protocadherin-beta.  
 GN PCDB20 OR PCDB14.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BAJB/C;  
 RA MEDLINE=2123055; PubMed=11322959;  
 RA Vanhalst K., Kools P., Vanden Eynde E., van Roy F.;  
 RT "The human and murine protocadherin-beta one-exon gene families show  
 RT high evolutionary conservation, despite the difference in gene  
 RT number."  
 RL FBS Lett. 495:120-125(2001).  
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.  
 DR EMBL: AF26313; AAK5328.1;  
 DR MGD: MGI:216758; Pcdhb20.  
 DR InterPro: IPR002106; AATRNA\_LigaseII.  
 DR InterPro: IPR002126; Cadherin.  
 DR Pfam: PF00028; cadherin; 5.  
 DR PROSITE: PS00179; AA\_TRNA\_Ligase\_II\_1; UNKNOWN\_1.  
 DR PROSITE: PS00232; CADHERIN\_1; UNKNOWN\_5.  
 DR PROSITE: PS00268; CADHERIN\_2; 5.  
 KM Calcium-binding; Cell adhesion; Glycoprotein.  
 SQ SEQUENCE 844 AA; 93034 MW; 303E780FF2E9E7C CRC64;

Query Match 70.1%; Score 2250; DB 11; Length 844;  
 Best local similarity 68.0%; Pred. No. 1.5e-147;  
 Matches 425; Conservative 94; Mismatches 106; Indels 0; Gaps 0;

QY 2 EAVVAHALOKROYSLICVFLGSMACAEPLRYVAEETRGFLANLAIDLGLVEELSA 61  
 Db 47 EGGGLSSCRKROYLVEFVGLGSAESLRYSAEETRGFLANLAIDLGLVEELSS 106  
 QY 62 RGRIVSDERTIGLLNPLTGLDLEKIDREELGPTPCVLPFOLLLEKFPQJFRAL 121  
 Db 107 REARVYSDDKKHLNLTGLDMLNRLDRELCSGTPCVLPFOVYLENPLQYRAEL 166

QY 122 WVRDINDSPVFLDREITLLESTPGATFELLESADSDVGINLRNTYSSNVFHLN 181  
 Db 167 HVRDINDSPFLDKETIKISTTGATFLENNQDDIGSNLQDYSISPSSTFYK 226  
 QY 182 VHDGEGNVSSELVLDKVLDRVEPELRLTLGLDGSPPRSCTTLIRLVLDINDNVE 241  
 Db 227 IHSDGDKITPELVLDALDHEESESELRLTLTALDGSPPRSCTTLIVKVIDINDNME 286  
 QY 242 FVESLIVQVQVENSPOGLVTVASARDLDTGSGEILVAFVTEERTLTFRINSQNL 301  
 Db 287 FAFSFTVQVPEMDPRLIOSSITLISADLDLMDGILGKISTYFLASEDIKRTFTFNPTSGEV 346  
 QY 302 HLAELNVEALIQYTTLTQAKDGGGSGKCTVYVHTDINDNPELMSLSPSPENSP 361  
 Db 347 NLRSLMDFEVIQSYVNIQATDGGGSACTLSVKTLDINDNAPVIMSSVKAIDEMS 406  
 QY 362 ETVVAVFRIDRSGNNAKAVCSIOHLPVLKPSVENEFTLVTERALDREERTENITIT 421  
 Db 407 ETVVAVFRIDRSGNNAKAVCSIOHLPVLKPSVENEFTLVTERALDREERTENITIT 466  
 QY 422 TVTDLGTPRLKTHNLTIVSDVNDNAPFTFSQTTTLRVRENNSPALHIGSVSATDRSG 481  
 Db 467 TASDKGTPLRTTQHTIRVOYSDINDNAPAFOTSTYTFYENNSPALHIGSVSATDRSG 526  
 QY 482 ANAQVYSLPPLPDPQLPGSLVSIADNGOLFALRSIDYEALEAFEFVGCATDRSGPAL 541  
 Db 527 SNAHITYSLPADPPELALASISINADNGOLFALRSIDYEALEAFEFVGCATDRSGPAL 586  
 QY 542 SSQALRVLVADANDNAPFLVPLONGSAPCTELVPRAEAGYLVAAVVDGSGQNAW 601  
 Db 587 SSQALRVLVADANDNAPFLVPLONGSAPCTELVPRAEAGYLVAAVVDGSGQNAW 646  
 QY 602 LSYQLKATEPGLFGVWANGEVRT 626  
 Db 647 LSYQLKATEPGLFGVWANGEVRT 671

## RESULT 6

Q96SE9 PRELIMINARY; PRT; 776 AA.  
 AC Q96SE9;  
 DT 01-DEC-2001 (TREMblrel. 19, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
 DE Protocadherin beta 8a.  
 GN PCDB8A.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=99308636; PubMed=10380929;  
 RA Wu Q., Maniatis T.;  
 RT "A striking organization of a large family of human neural cadherin-  
 RT live cell adhesion genes."  
 RL Cell 97:779-790(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20202599; PubMed=10716726;  
 RA Wu Q., Maniatis T.;  
 RT "Large exons encoding multiple ectodomains are a characteristic  
 RT feature of protocadherin genes."  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:3124-3129(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=21154914; PubMed=11230163;  
 RA Wu Q., Zhang T., Cheng J.-F., Kim Y., Grimwood J., Schmutz J.,  
 RA Dickson M., Noonan J.P., Zhang M.Q., Myers R.M., Maniatis T.;  
 RT "Comparative DNA sequence analysis of mouse and human protocadherin  
 RT gene clusters."  
 RL Genome Res. 11:389-404(2001).  
 CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.

DR EMBL: AY013878; AAK21988.1; -  
 DR InterPro: IPR002126; Cadherin.  
 DR Pfam: PF00028; cadherin.5.  
 DR PROSITE: PS00232; CADHERIN\_1; UNKNOWN\_5.  
 DR PROSITE: PS0268; CADHERIN\_2; 6.  
 DR PROSITE: PS0268; CADHERIN\_2; 6.  
 DR Calcium-binding; Cell adhesion; Glycoprotein.  
 DR Calcium-binding; Cell adhesion; Glycoprotein.  
 DR SEQUENCE 776 AA; 84983 MW; 03F420D3098D6032 CRC64;

Query Match 69.8%; Score 2240; DB 4; Length 776;  
 Best Local Similarity 69.2%; Pred. No. 6.6e-147;  
 Matches 434; Conservative 75; Mismatches 118; Indels 0; Gaps 0;

QY 1 MEARVHALQKQVSLVFLVGVSWAGAPLRYFAEETERTGFLANLAIDLGIVEELS 60  
 DB 1 MEIGMHNRRQVQLVFLVLLSLGAGELGYSVVEETEGSPFANLKGDLGLTMS 60  
 QY 61 ARGCIVSDETIGLLNPLTGDLLNKLDRRELCTGTPCVLPQQLKPRQIPFAE 120  
 DB 61 TKRARIISGKNQHLQAKQGTDDLINERLDRRELCTGTPCVLPQQLKPRQIPFAE 120  
 QY 121 LKVRDINDSPVFLDREITLNTLESTTGATFLSADSDVGINLKNYTISSNVYFI 180  
 DB 121 LKVRDINDSPVFLDREITLNTLESTTGATFLSADSDVGINLKNYTISSNVYFI 180  
 QY 181 NVHDNGEENVSELYLDKVLDRREVPBELRLTLTGDLGSSPPRSCTTLIRILVDINDVP 240  
 DB 181 LKHEFRDGRKPELVLDKELDRREVPBELRLTLTGDLGSSPPRSCTTLIRILVDINDVP 240  
 QY 241 EREVSELYKQVENSPLVYTVSAROLDGSGNCEIYAFYATERTLKTFRINSTGN 300  
 DB 241 EREVSELYKQVENSPLVYTVSAROLDGSGNCEIYAFYATERTLKTFRINSTGN 300  
 QY 301 LHKALNEALQIOTLTITQAKDGGGLSGCTVYVHTVDINDNPELLMSLTSPIENS 360  
 DB 301 VALRKQVDEMYTSEVRKATDGGGLSGCTVYVHTVDINDNPELLMSLTSPIENS 360  
 QY 361 PRTVAVFRIRDRSGNNKAVCSIODHLFVLPKPSVENEFTLVTERALDREERTYNT 420  
 DB 361 PEIVAVFVSDDPSGNNKTISSIOEDLPFLKPSVKNEYTLVTERALDREERTYNT 420  
 QY 421 ITVTDLGTRKLTQHNLTYSVDVNDNAPTESQTYTLRVRRNNSPALHIGSVATDRS 480  
 DB 421 LVTVDGTRKLTQHNLTYSVDVNDNAPTESQTYTLRVRRNNSPALHIGSVATDRS 480  
 QY 481 GANAQVTVSLRPPHDPOLPLGSLVSIADNQLFALRSIDFALQAFERRVAAADRGSPA 540  
 DB 481 GINAQVTVSLRPPHDPOLPLGSLVSIADNQLFALRSIDFALQAFERRVAAADRGSPA 540  
 QY 541 LSSQALVRLVADANDNAPFLVPLONGSAPCTELVPRAAEAGYLVAKVAVDGDGQNA 600  
 DB 541 LSSQALVRLVADANDNAPFLVPLONGSAPCTELVPRAAEAGYLVAKVAVDGDGQNA 600  
 QY 601 WLSTQALKATEPGLGVMAHNGEVRTA 627  
 DB 601 WLSTQALKATEPGLGVMAHNGEVRTA 627  
 RESULT 7  
 Q91Y00 PRELIMINARY; PRT; 799 AA.  
 AC Q91Y00;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Protocadherin beta 2.  
 GN PCDBH2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRNIN;

RX MEDLINE=99308636; PubMed=10380929;  
 RA Wu Q., Maniatis T.;  
 RT "A striking organization of a large family of human neural cadherin-  
 RT like cell adhesion genes.";  
 RT Cell 97:779-790(1999).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=BRNIN;  
 RX MEDLINE=20202599; PubMed=10716726;  
 RA Wu Q., Maniatis T.;  
 RT "Large exons encoding multiple ectodomains are a characteristic  
 RT feature of protocadherin genes.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 97:3124-3129(2000).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=BRNIN;  
 RX MEDLINE=21154914; PubMed=11230163;  
 RA Wu Q., Zhang T., Cheng J.-F., Kim Y., Grimwood J., Schmutz J.,  
 RA Dickson M., Noonan J.P., Zhang M.Q., Myers R.M., Maniatis T.;  
 RT "Comparative dna sequence analysis of mouse and human protocadherin  
 RT gene clusters.";  
 RT Genome Res. 11:389-404(2001).  
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.  
 DR EMBL: AY013781; AAK26070.1; -  
 DR MGD: MGI:2136735; Pcdhb2.  
 DR InterPro: IPR002126; Cadherin.  
 DR Pfam: PF00028; cadherin.5.  
 DR PROSITE: PS00232; CADHERIN\_1; UNKNOWN\_5.  
 DR PROSITE: PS0268; CADHERIN\_2; 5.  
 DR Calcium-binding; Cell adhesion; Glycoprotein.  
 DR Calcium-binding; Cell adhesion; Glycoprotein.  
 DR SEQUENCE 799 AA; 87528 MW; CAEB190E7AF967E CRC64;

Query Match 67.5%; Score 2167; DB 11; Length 799;  
 Best Local Similarity 68.0%; Pred. No. 8e-142;  
 Matches 427; Conservative 67; Mismatches 132; Indels 2; Gaps 1;

QY 1 MEA--RVVHALQKQVSLVFLVGVSWAGAPLRYFAEETERTGFLANLAIDLGIVEE 58  
 DB 1 MEAGEEMQILKRRQVLLFVLLGLIAQAGSTRIRYSEEBANENGVANLKLKGLVEEE 60  
 QY 59 LSAGCRVSDETIGLLNPLTGDLLNKLDRRELCTGTPCVLPQQLKPRQIPFAE 118  
 DB 61 LAAGPRVTSKGRKLNLEFNQGTDDLREKLDREELCTGTPCVLPQQLKPRQIPFAE 120  
 QY 119 AELVNDINDSPVFLDREITLNTLESTTGATFLSADSDVGINLKNYTISSNVYFI 178  
 DB 121 AELQRIQNDSPVFLDREITLNTLESTTGATFLSADSDVGINLKNYTISSNVYFI 180  
 QY 179 HINVDNGEENVSELYLDKVLDRREVPBELRLTLTGDLGSSPPRSCTTLIRILVDIND 238  
 DB 181 HLNQNSPDGTVLQVLVDKVLDRRESEIRLTLTGDLGSSPPRSCTTLIRILVDIND 240  
 QY 239 VPEVESLYKQVENSPLVYTVSAROLDGSGNCEIYAFYATERTLKTFRINST 298  
 DB 241 APVESKLYEVLQIPENSPLVYTVSAROLDGSGNCEIYAFYATERTLKTFRINST 300  
 QY 299 GNLHKAELNEALQIOTLTITQAKDGGGLSGCTVYVHTVDINDNPELLMSLTSPIE 358  
 DB 301 GEILLVKTLDPEISQATITICQATDGGGLSGSSVYIYQVADINDNPELLMSTFTNIPE 360  
 QY 359 NSPRTVAVFRIRDRSGNNKAVCSIODHLFVLPKPSVENEFTLVTERALDREERTYNT 418  
 DB 361 NAFETVAVFVSDDPSGNNKTISSIOEDLPFLKPSVKNEYTLVTERALDREERTYNT 420  
 QY 419 ITVTDLGTRKLTQHNLTYSVDVNDNAPTESQTYTLRVRRNNSPALHIGSVATDR 478  
 DB 421 ITIVSDGTRKLTQHNLTYSVDVNDNAPTESQTYTLRVRRNNSPALHIGSVATDR 480  
 QY 479 DSGANAQVTVSLRPPHDPOLPLGSLVSIADNQLFALRSIDFALQAFERRVAAADRGSPA 538  
 DB 481 DSGANAQVTVSLRPPHDPOLPLGSLVSIADNQLFALRSIDFALQAFERRVAAADRGSPA 540  
 QY 539 PALSSQALVRLVADANDNAPFLVPLONGSAPCTELVPRAAEAGYLVAKVAVDGDGQ 598









DR InterPro: IPR002126; Cadherin.  
 DR InterPro: IPR003580; Protachykinin.  
 DR Pfam: PF00028; cadherin. 5.  
 DR SMART: SM00203; TK; 1.  
 DR PROSITE: PS00232; CADHERIN\_1; UNKNOWN\_5.  
 DR PROSITE: PS50268; CADHERIN\_2; 6.  
 DR Calcium-binding; Cell adhesion; Glycoprotein.  
 KW SEQUENCE 792 AA; 86897 MW; F7CA3262A8E1D66 CRC64;

Query Match 66.6%; Score 2138; DB 11; Length 792;  
 Best Local Similarity 66.6%; Pred. No. 81e-140;  
 Matches 417; Conservative 80; Mismatches 129; Indels 0; Gaps 0;

QY 1 MEARVVALQKROVSLCFLVGVNAGAEPLRYFAEETERTGFLANLAIDLGLGVEELS 60  
 DB 1 MAARSCVSRKQVLFLLGLGLAGSELGRYSVTEETERTGSLVANLADLGLGVEALA 60  
 QY 61 ARGCRVDEETGFLNPLTGDLLNFKDLREELGCTECVCPQOLLLEKPPQIRAE 120  
 DB 61 AKRRVVCDDKKHFLDLSHGDLITNEKDLREKLCGTEPCMLYFQILMDNPQIRAE 120  
 QY 121 LVRNDINSHSPVFLDREITLNTLESTPGATFLLESADSDVGGINNRYNTISSNYEHI 180  
 DB 121 LRILIDINSHSPVFLDREITLNTLESTPGATFLLESADSDVGGINNRYNTISSNYEHI 180  
 QY 181 NVHNGEGNVSSELYLQKVIDREVEPELRLTLGLDGGSPRSCTTIRILVIDINDNP 240  
 DB 181 FVHNSDEMTPELVLDALDMEQPESTLTLTALDGGSPRSCTTIRILVIDINDNP 240  
 QY 241 EFVESLYKVQVSPVSLVVTASARDLDTGSGEIVYAFVTEKTLKTFRINSNGN 300  
 DB 241 QPQELVEIQAPENSPGLVYIKYTGEDVSGVAELSYFEDASELIRATFQINPESGE 300  
 QY 301 LHLKAEINVEAIQVYTLTQAKDGGSGKCTVAVVHTDINDNPELMSLSLTPSPENS 360  
 DB 301 IITKALDYERIKSYKINQAVDGGGLSARCTVLRVLVDNDAPELIMSLNEVEENS 360  
 QY 361 PETVAVAFRIRDRSGNNNAKWCSTIOHLEPRVLPKPSVENEYTLVTEALDREERTENIT 420  
 DB 361 PETVAVAFRIRDRSGNNNAKWCSTIOHLEPRVLPKPSVENEYTLVTEALDREERTENIT 420  
 QY 421 ITVTDLGPRLKQHNLTVTVDVNDNAPTFSTYTLRYRENNSPALHTIGSATDRDS 480  
 DB 421 ITVTDLGPRLKQHNLTVTVDVNDNAPTFSTYTLRYRENNSPALHTIGSATDRDS 480  
 QY 481 GANQVYSLPPLPDLPLGSLVSINADNGQLFALRSLDFEALQAEFEVGAADRSPA 540  
 DB 481 GSNHITYSLPPLDQMLTSLVSINADNGQLFALRSLDFEALQAEFEVGAADRSPA 540  
 QY 541 LSSQALVRYLVADANDNAPVLPVLPIONGSACTEIVPRAEAGYLVAKYAVVGDGQNA 600  
 DB 541 LSSQALVRYLVADANDNAPVLPVLPIONGSACTEIVPRAEAGYLVAKYAVVGDGQNA 600  
 QY 601 WLSTQALKATEPGLFGVMAHNGEVRT 626  
 DB 601 WLSTQALKATEPGLFGVMAHNGEVRT 626  
 RESULT 15  
 QY1VD8 PRELIMINARY; PRT; 799 AA.  
 AC QY1VD8;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Protocadherin beta 17.  
 GN PCDB17.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE-BRAIN;  
 RX MEDLINE-99308636; PubMed-10380929;  
 RA Wu Q., Maniatis T.;  
 RT "A striking organization of a large family of human neural cadherin-  
 RT like cell adhesion genes.";  
 RL Cell 97:779-790(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RX MEDLINE-20202599; PubMed-10716726;  
 RA Wu Q., Maniatis T.;  
 RT "Large exons encoding multiple ectodomains are a characteristic  
 RT feature of protocadherin genes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:3124-3129(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RX MEDLINE-21154914; PubMed-11230163;  
 RA Wu Q., Zhang T., Cheng J.-F., Kim Y., Grimwood J., Schmutz J.,  
 RA Dickson M., Noonan J.P., Zhang M.O., Myers R.M., Maniatis T.;  
 RT "Comparative dna sequence analysis of mouse and human protocadherin  
 RT gene clusters.";  
 RL Genome Res. 11:369-404(2001).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-EYE, AND RETINA;  
 RA Strausberg R.;  
 RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.  
 DR EMBL: AY013778; AKR26067.1; -;  
 DR EMBL: BC017149; AAH17149.1; -;  
 DR MGI: 2136754; Pcdh17.  
 DR InterPro: IPR002126; Cadherin.  
 DR Pfam: PF00028; cadherin. 5.  
 DR PROSITE: PS00232; CADHERIN\_1; UNKNOWN\_4.  
 DR PROSITE: PS50268; CADHERIN\_2; 6.  
 DR Calcium-binding; Cell adhesion; Glycoprotein.  
 KW SEQUENCE 799 AA; 87743 MW; 13807F0927972952 CRC64;

Query Match 65.6%; Score 2104; DB 11; Length 799;  
 Best Local Similarity 65.9%; Pred. No. 1.9e-137;  
 Matches 406; Conservative 93; Mismatches 127; Indels 0; Gaps 0;

QY 1 MEARVVALQKROVSLCFLVGVNAGAEPLRYFAEETERTGFLANLAIDLGLGVEELS 60  
 DB 1 MEAWMNCNROKOVAFVLLHVSAGAEIPGYSEETERTGFLANLGLVDLAELS 60  
 QY 61 ARGCRVDEETGFLNPLTGDLLNFKDLREELGCTECVCPQOLLLEKPPQIRAE 120  
 DB 61 NRRARIISQENKEHLQMLQSGDILLNFKDLREELGCTEPCVHLFVLMENPLEVQAE 120  
 QY 121 LVRNDINSHSPVFLDREITLNTLESTPGATFLLESADSDVGGINNRYNTISSNYEHI 180  
 DB 121 LVRNDINSHSPVFLDREITLNTLESTPGATFLLESADSDVGGINNRYNTISSNYEHI 180  
 QY 181 NVHNGEGNVSSELYLQKVIDREVEPELRLTLGLDGGSPRSCTTIRILVIDINDNP 240  
 DB 181 VTRNRSDEKRYPELVLEKELDREPEPELRLTLGLDGGAPRSGTAQVLLIEVDINDNP 240  
 QY 241 EFVESLYKVQVSPVSLVVTASARDLDTGSGEIVYAFVTEKTLKTFRINSNGN 300  
 DB 241 KFOQPTVQVQIPENSPGSLVTVSANDLSDGKVLALUSQSEDSITLLENVPTGE 300  
 QY 301 LHLKAEINVEAIQVYTLTQAKDGGSGKCTVAVVHTDINDNPELMSLSLTPSPENS 360  
 DB 301 IRLKEVDFERTIPSEVNDIKATDGGGLSGKCTLLKAVVDVNDNAPVMLALISLPVENS 360  
 QY 361 PETVAVAFRIRDRSGNNNAKWCSTIOHLEPRVLPKPSVENEYTLVTEALDREERTENIT 420  
 DB 361 PDEVVAVAVSPDPSANNGKMAISIEDLPELTKSSKNFYTLVTKRALDREERTENIT 420  
 QY 421 ITVTDLGPRLKQHNLTVTVDVNDNAPTFSTYTLRYRENNSPALHTIGSATDRDS 480  
 DB 421 ITVTDLGPRLKQHNLTVTVDVNDNAPTFSTYTLRYRENNSPALHTIGSATDRDS 480



```

Db 421 ITVSDLGTPRLTQHTITVQVSDTNDNAPAFNQTSTLFEVRENNSPAMHIGTISATDSDA 480
OY 481 CANAQVYTSLLPPIHPDPLPGSLVYSINADNGOLFALRSLDPEALQAFEFYRGADRGSPA 540
    |:::||||| |||| | ||:||||| |||||:::||||| ||| |||||
Db 481 GSNHSISYSLPSHDPOLADSLISINADNGOLFALRALDYEALQAFEFHVGAIIDGSPA 540
OY 541 LSSQALVRLVADANDNAPFVLYPLONGSAPCTELVPRAEAGYLVAKVAVADSDGSONA 600
    ||||||::: | ||||| |||||::: ||||| ||||| |||||
Db 541 LSSQALVRLVLDNDNAPFVLYPQNSSAPCTELLPRAAEPGYLVTKVAVADRDGSONA 600
OY 601 WLSYQLLKATEPGLFGVVAHNGEVRT 626
    ||:||||| ||||| ||||| |||||
Db 601 WLSFQLLKATEPGLFGVVAHNGEVRT 626

```

Search completed: May 16, 2003, 10:48:22  
 Job time : 41 secs



Qy	1	MEVVYHALOKROYSLCYEGLSVNAGAEPLRYPAEETEGCPALAIADIGLGYEELS	60
Db	1	MEAVVYHALOKROYSLCYEGLSVNAGAEPLRYPAEETEGCPALAIADIGLGYEELS	60
Qy	61	ARGCRIVSDETIGFLLNPLNLTGDLINKLDBREELCGTBEVCVLPFOLLERKPFQIRAE	12
Db	61	ARGCRIVSDETIGFLLNPLNLTGDLINKLDBREELCGTBEVCVLPOLLERKPFQIRAE	12
Qy	121	LWAVDINDHSPVLPDBREITLNLITLSTTPGATFLLLESASHDVGINNLRNVTISSNVYFHI	18
Db	121	LWAVDINDHSPVLPDBREITLNLITLSTTPGATFLLLESASHDVGINNLRNVTISSNVYFHI	18

QY 181 NVHNDGEGNVSELYLDKVLDEEVEPELRLTLTGLDGGSPRSGTTLRIILVLDINDVP 240  
DB 181 NVHNDGEGNVSELYLDKVLDEEVEPELRLTLTGLDGGSPRSGTTLRIILVLDINDVP 240  
QY 241 EEVESLYKVQVENSPPVSLVVTYSARDLDTGSGEIIYAFYATEFTLKTFRINSTGN 300  
DB 241 EEVESLYKVQVENSPPVSLVVTYSARDLDTGSGEIIYAFYATEFTLKTFRINSTGN 300  
QY 301 LHKALVYEAIOYITLTIOAKDGGSLGKCTVVVHTDINDNPELIMSSLTSPIDENS 360  
DB 301 LHKALVYEAIOYITLTIOAKDGGSLGKCTVVVHTDINDNPELIMSSLTSPIDENS 360  
QY 361 PETVAVERIRDRDGGNAKWKCSIOHLPVLKPSVNEFTVTERADDEEETENIT 420  
DB 361 PETVAVERIRDRDGGNAKWKCSIOHLPVLKPSVNEFTVTERADDEEETENIT 420  
QY 421 ITVTDLGTPLKTOHNLVTVDVNDNAPEFSQTTTLKRENNSPALHIGSVATDRDS 480  
DB 421 ITVTDLGTPLKTOHNLVTVDVNDNAPEFSQTTTLKRENNSPALHIGSVATDRDS 480  
QY 481 GANAQVYTSLLPDPHDPOLPGSLVSTINDNGOLFALNSLDFEALQAEFRVGAADRSPA 540  
DB 481 GANAQVYTSLLPDPHDPOLPGSLVSTINDNGOLFALNSLDFEALQAEFRVGAADRSPA 540  
QY 541 LSSQALVYLVADANDNAPFVLYPLONGSAPCTELVPRAAGVLVAKVAVDGDSCQA 600  
DB 541 LSSQALVYLVADANDNAPFVLYPLONGSAPCTELVPRAAGVLVAKVAVDGDSCQA 600  
QY 601 WLSTYQLKATEPGLGFWAHNGEVRTA 627  
DB 601 WLSTYQLKATEPGLGFWAHNGEVRTA 627

RESULT 2  
US-10-036-041-52  
Sequence 52, Application US/10036041  
Publication No. US20020192751A1  
GENERAL INFORMATION:  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3030R1C8  
CURRENT APPLICATION NUMBER: US/10/036, 041  
PRIOR FILING DATE: 2001-12-26  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/112514  
PRIOR FILING DATE: 1998-12-15  
PRIOR APPLICATION NUMBER: 60/113300  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: 60/113430  
PRIOR FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: 60/113605  
PRIOR FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: 60/113621  
PRIOR FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: 60/114140  
PRIOR FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: 60/115552  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/116843  
PRIOR FILING DATE: 1999-01-22  
PRIOR APPLICATION NUMBER: 60/125774  
PRIOR FILING DATE: 1999-03-23

PRIOR APPLICATION NUMBER: 60/125778  
PRIOR FILING DATE: 1999-03-23  
PRIOR APPLICATION NUMBER: 60/125826  
PRIOR FILING DATE: 1999-03-24  
PRIOR APPLICATION NUMBER: 60/127035  
PRIOR FILING DATE: 1999-03-31  
PRIOR APPLICATION NUMBER: 60/127706  
PRIOR FILING DATE: 1999-04-05  
PRIOR APPLICATION NUMBER: 60/129122  
PRIOR FILING DATE: 1999-04-13  
PRIOR APPLICATION NUMBER: 60/130359  
PRIOR FILING DATE: 1999-04-21  
PRIOR APPLICATION NUMBER: 60/131270  
PRIOR FILING DATE: 1999-04-27  
PRIOR APPLICATION NUMBER: 60/131272  
PRIOR FILING DATE: 1999-04-27  
PRIOR APPLICATION NUMBER: 60/131291  
PRIOR FILING DATE: 1999-04-27  
PRIOR APPLICATION NUMBER: 60/132371  
PRIOR FILING DATE: 1999-05-04  
PRIOR APPLICATION NUMBER: 60/132379  
PRIOR FILING DATE: 1999-05-04  
PRIOR APPLICATION NUMBER: 60/132383  
PRIOR FILING DATE: 1999-05-04  
PRIOR APPLICATION NUMBER: 60/135750  
PRIOR FILING DATE: 1999-05-25  
PRIOR APPLICATION NUMBER: 60/138166  
PRIOR FILING DATE: 1999-06-08  
PRIOR APPLICATION NUMBER: 60/144791  
PRIOR FILING DATE: 1999-07-20  
PRIOR APPLICATION NUMBER: 60/146970  
PRIOR FILING DATE: 1999-08-03  
PRIOR APPLICATION NUMBER: 60/162506  
PRIOR FILING DATE: 1999-10-29  
PRIOR APPLICATION NUMBER: 09/311832  
PRIOR FILING DATE: 1999-05-14  
PRIOR APPLICATION NUMBER: 09/380142  
PRIOR FILING DATE: 1999-08-25  
PRIOR APPLICATION NUMBER: 09/644848  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: 09/747259  
PRIOR FILING DATE: 2000-12-20  
PRIOR APPLICATION NUMBER: 09/816744  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: 09/854208  
PRIOR FILING DATE: 2001-05-10  
PRIOR APPLICATION NUMBER: 09/854280  
PRIOR FILING DATE: 2001-05-10  
PRIOR APPLICATION NUMBER: 09/874503  
PRIOR FILING DATE: 2001-06-05  
PRIOR APPLICATION NUMBER: 09/869599  
PRIOR FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: 09/908, 827  
PRIOR FILING DATE: 2001-07-18  
PRIOR APPLICATION NUMBER: PCT/US99/10733  
PRIOR FILING DATE: 1999-05-14  
PRIOR APPLICATION NUMBER: PCT/US99/28551  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30720  
PRIOR FILING DATE: 1999-12-22  
PRIOR APPLICATION NUMBER: PCT/US00/05601  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: PCT/US00/05841  
PRIOR FILING DATE: 2000-03-02  
PRIOR APPLICATION NUMBER: PCT/US00/14042  
PRIOR FILING DATE: 2000-05-22  
PRIOR APPLICATION NUMBER: PCT/US00/15264  
PRIOR FILING DATE: 2000-06-02  
PRIOR APPLICATION NUMBER: PCT/US00/23522  
PRIOR FILING DATE: 2000-08-23  
PRIOR APPLICATION NUMBER: PCT/US00/23328  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: PCT/US00/32678

;; PRIOR FILING DATE: 2000-12-01  
;; PRIOR APPLICATION NUMBER: PCT/US00/34956  
;; PRIOR FILING DATE: 2000-12-20  
;; PRIOR APPLICATION NUMBER: PCT/US01/06520  
;; PRIOR FILING DATE: 2001-02-28  
;; PRIOR APPLICATION NUMBER: PCT/US01/17800  
;; PRIOR FILING DATE: 2001-06-01  
;; PRIOR APPLICATION NUMBER: PCT/US01/19692  
;; PRIOR FILING DATE: 2001-06-20  
;; PRIOR APPLICATION NUMBER: PCT/US01/21066  
;; PRIOR FILING DATE: 2001-06-29  
;; PRIOR APPLICATION NUMBER: PCT/US01/21735  
;; PRIOR FILING DATE: 2001-07-09  
;; NUMBER OF SEQ ID NOS: 80  
;; SEQ ID NO 52  
;; LENGTH: 800  
;; TYPE: PRT  
;; ORGANISM: Homo Sapien  
US-10-036-041-52

Query Match Best Local Similarity 70.1%; Score 2249; DB 9; Length 800;

Matches 435; Conservative 70; Mismatches 113; Indels 0; Gaps 0;

QY 10 QKROYSLCVPLGYSMAGAEPLRYFAEEERGTFLANLAIDGLGYEELASRGCRIVSD 69  
Db 10 RORVLFLEFMVGSLSGSGRVSVEEPEKGSFVNNAKDLGLAGELARCTRVSD 69  
QY 70 ETIGFLINPLITGDLLEKIDREELGPTPECVLPQOLLEKFPQIFRAELWVDINDH 129  
Db 70 DNKQYLLDSTGNLNEKIDREELGPTPECVLPQOLLEKFPQIFRAELWVDINDH 129  
QY 130 SPVFLREITNLESTPGATFLLESADSDVGINNANTISSYNNVFNHNDGECN 189  
Db 130 APVQDKETVYKISENTAEGTAFLERARQDPDGLNGIQTITSPNSFHNIGSGEGM 189  
QY 190 VYSELVLDKVLDRREPELRLTLGLDGGSPRSQTILRLVLDINDNPEVESLYKV 249  
Db 190 IYPELVLDKALDREBELSLTLVLDGGSPRSQTSTVRIVLDVNDNAQFQALYET 249  
QY 250 QVPEKSPVGSIVYVSADLDTGSGELVYAFYATEKTKTFRKINSTGMLKAEIN 309  
Db 250 QAPENSPIGFLIVKWMADVDGVAEVSYSFDPASENIRTPQINPFSGEIFLRELDY 309  
QY 310 EAIQYTLTIOAKDGGGSGCTVYVHTDINDNPELLMSLSPIPENSPEIVAVFR 369  
Db 310 ELVNSYKINIOAMDGGGSLASRCRVLYEVLINDNPELLVSSFSNSVAENSPETPLAVFK 369  
QY 370 IRDRDSGNNAKWCSTODHLEPVLPKPSYENFTLVTERALDRERTEYNTITVTDLGP 429  
Db 370 INDRDSGNGKMWCVYIQENLEPLLPKPSYENFTLITBGLADREIRAEVNTITVTDLGP 429  
QY 430 RLKTOHNLTVYSDVNDNAPFTSQTYYTLRKRENNSPALHIGSYSADRDGANAQVYTS 489  
Db 430 RLKTEHNTLVYSDVNDNAPFTSQTYYTLRKRENNSPALHIGSYSADRDGANAQVYTS 489  
QY 490 LLPHPDPLGSLVYSINADNGOLFALRSILDFEALQAEFFVGAADGSPALSSQALVRY 549  
Db 490 LLPHPDPLPLASLVSINADNGHLPALRSILDEALQAEFFVGAADGSPALSSQALVRY 549  
QY 550 LVADANDAPFVLYPLONGSAPCTELVPRAAEAGLVAKVYAVVGDGSGONAMLSYOLKA 609  
Db 550 LVLDANDNSPFLVPLONGSAPCTELVPRAAEAGLVAKVYAVVGDGSGONAMLSYOLKA 609  
QY 610 TEPGLFGVWANGEVYTA 627  
Db 610 TEPGLFGVWANGEVYTA 627

RESULT 3  
US-10-035-855-52  
Sequence 52, Application US/10035855  
Publication No. US20030008348A1

;; GENERAL INFORMATION:  
;; APPLICANT: Desnoyers, Luc  
;; APPLICANT: Eaton, Dan L.  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul J.  
;; APPLICANT: Gurney, Austin L.  
;; APPLICANT: Pan, James  
;; APPLICANT: Stewart, Timothy A.  
;; APPLICANT: Watanabe, Colin K.  
;; APPLICANT: Wood, William I.  
;; APPLICANT: Zhang, Demin  
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
;; FILE REFERENCE: P3030RICA  
;; CURRENT APPLICATION NUMBER: US/10/035,855  
;; CURRENT FILING DATE: 2001-12-26  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/112514  
;; PRIOR FILING DATE: 1998-12-15  
;; PRIOR APPLICATION NUMBER: 60/113300  
;; PRIOR FILING DATE: 1998-12-22  
;; PRIOR APPLICATION NUMBER: 60/113430  
;; PRIOR FILING DATE: 1998-12-23  
;; PRIOR APPLICATION NUMBER: 60/113605  
;; PRIOR FILING DATE: 1998-12-23  
;; PRIOR APPLICATION NUMBER: 60/113621  
;; PRIOR FILING DATE: 1998-12-23  
;; PRIOR APPLICATION NUMBER: 60/114140  
;; PRIOR FILING DATE: 1998-12-23  
;; PRIOR APPLICATION NUMBER: 60/115552  
;; PRIOR FILING DATE: 1999-01-12  
;; PRIOR APPLICATION NUMBER: 60/116843  
;; PRIOR FILING DATE: 1999-01-22  
;; PRIOR APPLICATION NUMBER: 60/125774  
;; PRIOR FILING DATE: 1999-03-23  
;; PRIOR APPLICATION NUMBER: 60/125778  
;; PRIOR FILING DATE: 1999-03-23  
;; PRIOR APPLICATION NUMBER: 60/125826  
;; PRIOR FILING DATE: 1999-03-24  
;; PRIOR APPLICATION NUMBER: 60/127035  
;; PRIOR FILING DATE: 1999-03-31  
;; PRIOR APPLICATION NUMBER: 60/12706  
;; PRIOR FILING DATE: 1999-04-05  
;; PRIOR APPLICATION NUMBER: 60/129122  
;; PRIOR FILING DATE: 1999-04-13  
;; PRIOR APPLICATION NUMBER: 60/130359  
;; PRIOR FILING DATE: 1999-04-21  
;; PRIOR APPLICATION NUMBER: 60/131270  
;; PRIOR FILING DATE: 1999-04-27  
;; PRIOR APPLICATION NUMBER: 60/131272  
;; PRIOR FILING DATE: 1999-04-27  
;; PRIOR APPLICATION NUMBER: 60/131291  
;; PRIOR FILING DATE: 1999-04-27  
;; PRIOR APPLICATION NUMBER: 60/132371  
;; PRIOR FILING DATE: 1999-05-04  
;; PRIOR APPLICATION NUMBER: 60/132379  
;; PRIOR FILING DATE: 1999-05-04  
;; PRIOR APPLICATION NUMBER: 60/132383  
;; PRIOR FILING DATE: 1999-05-04  
;; PRIOR APPLICATION NUMBER: 60/135750  
;; PRIOR FILING DATE: 1999-05-25  
;; PRIOR APPLICATION NUMBER: 60/138166  
;; PRIOR FILING DATE: 1999-06-08  
;; PRIOR APPLICATION NUMBER: 60/144791  
;; PRIOR FILING DATE: 1999-07-20  
;; PRIOR APPLICATION NUMBER: 60/146970  
;; PRIOR FILING DATE: 1999-08-03  
;; PRIOR APPLICATION NUMBER: 60/162506  
;; PRIOR FILING DATE: 1999-10-29  
;; PRIOR APPLICATION NUMBER: 09/311832  
;; PRIOR FILING DATE: 1999-05-14  
;; PRIOR APPLICATION NUMBER: 09/380142

```

PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/644848
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 09/747259
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 09/816744
PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 09/854208
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/854280
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/874503
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: 09/869599
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/908, 827
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: PCT/US99/10733
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30720
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: PCT/US00/05601
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/23522
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: PCT/US00/34956
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: PCT/US01/17800
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: PCT/US01/19692
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: PCT/US01/22106
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: PCT/US01/21735
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 80
SEQ ID NO 52
LENGTH: 800
TYPE: PRT
ORGANISM: Homo Sapien
US-10-035-855-52

```

Query Match 70.1%; Score 2249; DB 9; Length 800;  
 Best Local Similarity 70.4%; Pred. No. 1.2e-157;  
 Matches 435; Conservative 70; Mismatches 113; Indels 0; Gaps 0;

```

QY 10 OKROYSLICVFLGVSAGAEPLRTFAEETRGFTFLANLAIIDLGVEELISARGCRIVSD 69
DB 10 RQROYLFLFVWGSVLAGSGFRYSVTEETKSGFVNVNLAIDGLAEGELANGSTRVSD 69
QY 70 ETIGFLLINPLTGDLINLEKLDREELCGPTEPCVLPOLLLEKFPQIFRAELRWINDH 129
DB 70 DNKOYLLDLSHTGNLNLNEKLDREELCGPTEPCVLPOLLLEKFPQIFRAELRWINDH 129
QY 130 SPVFLDEITLNLLESTTPGATELLBSAHDSDVGINNLRYTSSNVYFHINVDGEGN 189
DB 130 APVPODEYVLTAKISENTAECTAFRLERAQDPDGLINCIOVTTSPNSFHHINISGDEGN 189
QY 190 VYSELVLDKVLDEEYVELRLTLTGIDGSPRSSGFTLLIRIILVDINDNVPEVESLKYV 249

```

```

DB 190 IYPELVLDKALDREOQELSLTILALDGGSPSSGTSVRIIVLDVNDNAPQAPQALYER 249
QY 250 QVPSNPVSGSLVYVYSARDLDTGNGCIYVAFYATERTLKTIRINSGNLHLKALYN 309
DB 250 QAPNSPIGLVYKVMADVDGVAEVSYSFFDASINRTTQIINFSGEILFRLRLDY 309
QY 310 EAIOTYTLTIOAKDGGLSGKCTVWVHTDINDNPELLMSLTSPIPEPSPEVAVR 369
DB 310 ELVYSKINIQAADGGLSARCRVIVEVDTONNPELLIVSSNSVAENSPETPLAVRK 369
QY 370 IRRDSSNNAKMVCISIDHLPFLVKSVEVFYLVTERALDREREYNTITVTDLGRP 429
DB 370 INRDSGNGKWCYIOENLPLFKPSVENFYILITEGALDREIRAEYNTITVTDLGRP 429
QY 430 RLKQHNLTVTVSDVNDNAPFTSQTYTLRENNSPALHIGSVSTDDSDGNAQVYS 489
DB 430 RLKTEHNTVTVSDVNDNAPFTSQTYTLRENNSPALHIGSVSTDDSDGNAQVYS 489
QY 490 LPPHDFPLGSLVSTINADNGQLFALRSIDFALQAFERVGADRGSPALSSQALVRY 549
DB 490 LPPHDFPLGSLVSTINADNGQLFALRSIDFALQAFERVGADRGSPALSSQALVRY 549
QY 550 LVADANDNAPFVLYPLONGSAPCTELVPRAAEGYVAAVAVDGGSGONAMLSYOLKA 609
DB 550 LVADANDNAPFVLYPLONGSAPCTELVPRAAEGYVAAVAVDGGSGONAMLSYOLKA 609
QY 610 TERGLFGVMAHNGEVRTA 627
DB 610 TERGLFGVMAHNGEVRTA 627

```

RESULT 4  
 US-10-174-590-478  
 Sequence 478, Application US/10174590  
 Publication No. US20030008352A1  
 GENERAL INFORMATION:  
 APPLICANT: Baker, Kevin P.  
 APPLICANT: Chen, Jian  
 APPLICANT: Desnoyers, Luc  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Godowski, Paul J.  
 APPLICANT: Gurney, Austin L.  
 APPLICANT: Pan, James  
 APPLICANT: Smith, Victoria  
 APPLICANT: Watanabe, Colin K.  
 APPLICANT: Wood, William I.  
 APPLICANT: Zhang, Zemin  
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 FILE REFERENCE: P3430R1C42  
 CURRENT APPLICATION NUMBER: US/10/174,590  
 PRIOR FILING DATE: 2002-06-18  
 PRIOR application removed - See file wrapper or Palm  
 NUMBER OF SEQ ID NOS: 612  
 SEQ ID NO 478  
 LENGTH: 800  
 TYPE: PRT  
 ORGANISM: Homo Sapien  
 US-10-174-590-478

Query Match 70.1%; Score 2249; DB 9; Length 800;  
 Best Local Similarity 70.4%; Pred. No. 1.2e-157;  
 Matches 435; Conservative 70; Mismatches 113; Indels 0; Gaps 0;

```

QY 10 OKROYSLICVFLGVSAGAEPLRTFAEETRGFTFLANLAIIDLGVEELISARGCRIVSD 69
DB 10 RQROYLFLFVWGSVLAGSGFRYSVTEETKSGFVNVNLAIDGLAEGELANGSTRVSD 69
QY 70 ETIGFLLINPLTGDLINLEKLDREELCGPTEPCVLPOLLLEKFPQIFRAELRWINDH 129
DB 70 DNKOYLLDLSHTGNLNLNEKLDREELCGPTEPCVLPOLLLEKFPQIFRAELRWINDH 129

```

	Query Match	70.1%	Score 2249;	DB 97	length 800;
	Best Local Similarity	70.4%	Pred. No. 1,2e-157;		
	Matches 435;	Conservative 70;	Mismatches 113;	Indels 0;	Gaps 0;
QY	10 QKROYSLICVFLGYSNAGCEPIRYFAETETGTFPLANLAIDLGCEVLSRGRIVSD	69			
DB	10 RROVAFLEFLMGVSLAGSGFGNRYSTETETGSGFPVNIAKIDJLAEGELGARGIVSD	69			

Query Match	70.1%;	Score 2249;	DB 9;	Length 800;
Best Local Similarity	70.4%;	Pred. No. 1.2e-157;		

	Matches	435; Conservative	70; Mismatches	113; Indels	0; Gaps
QY	10	OKROYSLICVFLIGVSWAGAEPLRYFAAEETEGTFLIANLAIDLGLVEELSLARCCIVSD	69		
Db	10	RORVLLFLFLFGLVGVSLAGSGFGRYSTTEETEGSFVYNAIAKDLGAEGELARAGTRVSD	69		
QY	70	ETIGELLNPLNGDILLNKLREBELCGTPEECVLPFOILLKEKFPQIPAEIWMVRINDH	129		
Db	70	DNKOYLLDSDSHNGLLTNKLDREKLCKGKECMILYFOILMDPQIYIAELFLVRINDH	129		
QY	130	SVFVLDREITLLILESTTFGATFLLBSAHDSDVGINNENRNTYSSNVYFHINVDGEGN	189		
Db	130	AVFEDKKEFLVLIKSENTAGTAFRLERADDPGGINGIONNTISPNSFFHINISGDEGN	189		
QY	190	VYSEELVDKVLDREREVELRLTLTGIDGSPRSRGTTLRIILVILINCVPEFVSLVY	249		
Db	190	IYPELVLDKALDREGEGLSLTLTALMDGSPRSRGTSTYRIYVLDVNDNAPQFAQHLVET	249		
QY	250	QVPENSPVGSLVYVYASARDLTGSGEIVYAFVYATERTLKTFRINSTSGNLHLAEALNY	309		
Db	250	QAPENSPGFLFLVKWMAEDVDGVAEAESYSFEDASENIRTFQINPESGEIFLRELIDY	309		
QY	310	EAIQVYTLTLQAKDGGSGKCTVYVHVHTDINDNPEILMSLSTPIPNSEVYVAVR	369		
Db	310	EIVNSYKINIQAMDGGSLARCRVLYEVLIDNDNPELIVSSFSVANSSETPLAVFK	369		
QY	370	IRDRSGNNAKWCSIQDHLPEVLKPSYENFTVLTERRALDREERTENITTTVDLGP	429		
Db	370	INDRSGENGKAWCYIOENLPELKPSENFYILITEGDLREIAEYNTITTVDLGP	429		
QY	430	RLKTHNLTVMYSDVDNDNAPTFSSQTYTLTLVRENNPSALHIGSVATDRDSANNOVYS	489		
Db	430	RLKTEHNITLISDVNDNAPAFOTSYTLTLFRENNPSALHIGSVATDRDSANNOVYS	489		
QY	490	LLEPHDPLPLGSLVLSINADNGQLFALSDFEALQAEFFRVGADRDSPALSSQALRV	549		
Db	490	LLPPQDPHLPLASLVSINADNGHLFALUSDLEALQAEFFRVGATDRKSPALSRALRV	549		
QY	550	LYAANDNAPFVLYPLONGSAPCTBELVYRAADAGLYAKVAVVDGSDQAMWLSYQLLKA	609		
Db	550	LYLANDNSPVLVPLONGSAPCTELVYRAAEPGLVYKVAVVDGSDQAMWLSYQLLKA	609		
QY	610	TEPGLFGVWANGVEYRTA 627			
Db	610	TEPGLFGVWANGVEYRTA 627			
RESULT 7					
US-10-173-706-478					
Sequence 478, Application US/10173706					
Publication No. US20030022293A1					
GENERAL INFORMATION:					
APPLICANT: Baker, Kevin P.					
APPLICANT: Chen, Jian					
APPLICANT: Desnoyers, Luc					
APPLICANT: Goddard, Audrey					
APPLICANT: Godowski, Paul J.					
APPLICANT: Gurney, Austin L.					
APPLICANT: Pan, James					
APPLICANT: Smith, Victoria					
APPLICANT: Matanabe, Colin K.					
APPLICANT: Wood, William I.					
APPLICANT: Zhang, Zemin					
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC					
FILE OF INVENTION: ACIDS ENCODING THE SAME					
FILE REFERENCE: P3430R1C7					
CURRENT APPLICATION NUMBER: US/10/173,706					
CURRENT FILING DATE: 2002-06-17					
Prior Application removed - See file Wrapper or Palm					
NUMBER OF SEQ ID NOS: 612					
SEQ ID NO 478					
LENGTH: 800					
TYPE: PRT					

```

; ORGANISM: Homo Sapien
US-10-173-706-478

Query Match          70.1%, Score 2249; DB 9; Length 800;
Best Local Similarity 70.4%, Pred. No. 1.2e-15;
Matches 435; Conservative 70; Mismatches 113; Indels 0; Gaps 0;

QY      10 QKRVSLICVFEFVGMAAEPFLRYFAVEETEGTFLANLIDLGLGVEELSAARGRIYS D 69
       ::::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      10 RQRQLFLFLFWGVSLASGFGFRYSVTBETEKGSTVVNLAKDLGLAEGLAARGRIVSD   69
QY      70 ETIGELLNPLTGDLLNEKLDRBELCGPTEPCVLFPOLLLEKPPQIFRAELWBDINDH 129
       : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      70 DNKRQILDLDSHGNTLTETREKLDREKLCGPKECKMYEFLIMDDPEQIYRAELRVIDNH 129
QY      130 SPVFEDRETTLINLESTPGATFLESADSPVGINLRNTYISSNVYFHIVHNGGN 189
       : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      130 APVFDKRTYALKISNTALDEGRAFLERKADDPDGGINGNTITSNPFHINISGDCGM 189
QY      190 VYSVELDKVIDREEVPELRLLTLTGLDGSGPPRSQTLLIRILLVDINDVPEVESLYKV 249
       : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      190 IYPELVLDKADIREOGELSTLTFTALDGGSSRSGSTSVRIYVLDVNNAPOFAQALYET 249
QY      250 QVPENSPVGSIVVMYASARDLDTGNSGEIVYAFFYATEERTLKTFRINTSGNHLKAELNY 309
       : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      250 QAPENSPGFELTVKAWADVDSGVAAEVSYSEFDASENIRTFQNPINSGETLFRELIDY 309
QY      310 EAIQTYTLTIQAOKDGGLSGRCQTVVHVHTJINDNPELMLSLTPISNSETVVAVR 369
       : : : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      310 ELVNSYKINIQAMDGGCSARCRLVLEYLDINDNPPELLIVSFSNSVANSEPFLAVFK 369
QY      370 IRDRPSGNNAKMGSIQDHLRPVLKPSVENYTYLTTERALDREETENVITYTDLCTP 429
       : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      370 INDRSGENGKMVCYIOENLPFLKPSVENFYILTTEALDRETRAEYNITYTYDICTP 429
QY      430 RLKTOHNTLVSDVNDNAPFESQTTYTLRVRENNSPALIHGISATSATRDGANNOVYS 489
       : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      430 RLKTEHNITTVSDVNDNAPAFOTSTYTLPFRENNSPALIHGISATSATRDGSTRNQVYS 489
QY      490 LLPHPDPLPLGSLVSIADNGQLFALRSLDPEALQAEFFRVGAADRGSFALLSCVALRV 549
       : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      490 LLPDPDPTPLASIVSIIADNGHFLALSRLDYEAQLQAEFFRGVATDRGPSALTREALRV 549
QY      550 LVAPANDAPAPVLPLONGSAPCTEIVPRAEAGIVAKVAVDGDSONMALSYYOLIKA 609
       : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      550 LVDANDNSPVVLPLONGSAPCTELVPRAEPGLIVTKVAVDGDSONAMLSYOLIKA 609
QY      610 TEGPLFGVMAHNGEVRTA 627
       : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      610 TEGPLFGVMAHNGEVRTA 627

RESULT 8
US-10-175-738-478
Sequence 478, Application US/10175738
Publication No. US20030022294A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME
FILE REFERENCE: P3430RIC45
CURRENT APPLICATION NUMBER: US/10/175, 738
CURRENT FILING DATE: 2002-06-19

```



; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Wood, William I.  
 ; APPLICANT: Zhang, Zemin

```

;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Gurney, Austin L.

```

```

; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C70
; CURRENT APPLICATION NUMBER: US/10/176,482
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 478
; LENGTH: 800
; TYPE: PRP
; ORGANISM: Homo Sapien
US-10-176-482-478

```

```

Query Match      70.1%; Score 2249; DB 9; Length 800;
Best Local Similarity 70.4%; Pred. NO. 1.2e-157;
Matches 435; Conservative 70; Mismatches 113; Indels 0; Gaps 0;

```

```

QY 10 OKROYSLLCVFLGVSMAGAEPLRYFAETERTGTFLANLADLGVEELSGRCRIYSD 69
Db 10 RORQVLFLEFLMGVSLAGSGFRYSVTEETEKSGFVNKLADGLAGELAAKGRIVSD 69
QY 70 ETIGFLNPLITGDLLEKLDREELCGPTEPCVLPOLLLEKPFQIRAEIWRVDINDH 129
Db 70 DNKOYLLDLSHTGNLLTNEKLDREELCGPTEPCVLPOLLLEKPFQIRAEIWRVDINDH 129
QY 130 SPVFLDREITLLESTTPGATFLLESADSDVGINNLRNTTSSNVFFHINVDNDEGN 189
Db 130 APVQDKETVLKISENTAGTAFLERADDPDGGNGIQNTYISPSNFHINISGDEGM 189
QY 190 VYSELVDKVDLREVEBELRLTLTGDLGSPRSSTLIRILYLDINDNPEFEVESLYKV 249
Db 190 IYPELVLDKALDREOGEELSLTLTALDGSRSRSGTSTVRILYLDVNDNAQFOALYET 249
QY 250 QVPENSPVGLVTVYSARDLDTGSGEIVYAFYATERITKTRINSTGNLHKAEIYN 309
Db 250 QAPENSPIGFLIVKVAEDVDGVAEVSYSFEDASENIRTFQINFSGEIFLRELDY 309
QY 310 EIAYOTYTLTIQAKDGGGLSGKCTVYVHTDINDNPELLMSLSPIPENSPETVAVFR 369
Db 310 ELVNSYKINTIQAMGGGLSARCRVLYEVLDTNDNPELLVSSFSNSVAENSPETPLAVFR 369
QY 370 IRODROGNNAKMVCISIDHLPVILKPSVENFTLYTERALDREERTENITITVTDLGP 429
Db 370 INDRDSENGKMKVCYIQENLPFLKPSVENFTLITGALDREIRAEVNTITVTDLGP 429
QY 430 RLKTOHNLTVYSDVNDNAPFTSQTYTLRVRENNSPALHIGVSATDRDSCANAQVYS 489
Db 430 RLKTEHNTIVLVSDVNDNAPFTSQTYTLRVRENNSPALHIGVSATDRDSCANAQVYS 489
QY 490 LLPHPDPLPGSLVSIINADNGQLFALRSIDFELQAEFFRVGAAGSPALSSQALVRY 549
Db 490 LLPHPDPLPLASVSIINADNGHLFALRSIDYELQAEFFRVGATBPGSPALSSREALVRY 549
QY 550 LVADANDNAPFVLYPLQNGSAPCTEIVPRAAEAGYLVAKVYAVDGGSGONAMLSYOLLA 609
Db 550 LVADANDNSPFVLYPLQNGSAPCTEIVPRAAEPGYLVTKVYAVDGGSGONAMLSYOLLA 609
QY 610 TERGLFGVMAHNGEYRTA 627
Db 610 TERGLFGVMAHNGEYRTA 627

```

```

RESULT 11
US-10-176-757-478
; Sequence 478, Application US/10176757
; Publication No. US20030022297A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

```

```

; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C86
; CURRENT APPLICATION NUMBER: US/10/176,757
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 478
; LENGTH: 800
; TYPE: PRP
; ORGANISM: Homo Sapien
US-10-176-757-478

```

```

Query Match      70.1%; Score 2249; DB 9; Length 800;
Best Local Similarity 70.4%; Pred. NO. 1.2e-157;
Matches 435; Conservative 70; Mismatches 113; Indels 0; Gaps 0;

```

```

QY 10 OKROYSLLCVFLGVSMAGAEPLRYFAETERTGTFLANLADLGVEELSGRCRIYSD 69
Db 10 RORQVLFLEFLMGVSLAGSGFRYSVTEETEKSGFVNKLADGLAGELAAKGRIVSD 69
QY 70 ETIGFLNPLITGDLLEKLDREELCGPTEPCVLPOLLLEKPFQIRAEIWRVDINDH 129
Db 70 DNKOYLLDLSHTGNLLTNEKLDREELCGPTEPCVLPOLLLEKPFQIRAEIWRVDINDH 129
QY 130 SPVFLDREITLLESTTPGATFLLESADSDVGINNLRNTTSSNVFFHINVDNDEGN 189
Db 130 APVQDKETVLKISENTAGTAFLERADDPDGGNGIQNTYISPSNFHINISGDEGM 189
QY 190 VYSELVDKVDLREVEBELRLTLTGDLGSPRSSTLIRILYLDINDNPEFEVESLYKV 249
Db 190 IYPELVLDKALDREOGEELSLTLTALDGSRSRSGTSTVRILYLDVNDNAQFOALYET 249
QY 250 QVPENSPVGLVTVYSARDLDTGSGEIVYAFYATERITKTRINSTGNLHKAEIYN 309
Db 250 QAPENSPIGFLIVKVAEDVDGVAEVSYSFEDASENIRTFQINFSGEIFLRELDY 309
QY 310 EIAYOTYTLTIQAKDGGGLSGKCTVYVHTDINDNPELLMSLSPIPENSPETVAVFR 369
Db 310 ELVNSYKINTIQAMGGGLSARCRVLYEVLDTNDNPELLVSSFSNSVAENSPETPLAVFR 369
QY 370 IRODROGNNAKMVCISIDHLPVILKPSVENFTLYTERALDREERTENITITVTDLGP 429
Db 370 INDRDSENGKMKVCYIQENLPFLKPSVENFTLITGALDREIRAEVNTITVTDLGP 429
QY 430 RLKTOHNLTVYSDVNDNAPFTSQTYTLRVRENNSPALHIGVSATDRDSCANAQVYS 489
Db 430 RLKTEHNTIVLVSDVNDNAPFTSQTYTLRVRENNSPALHIGVSATDRDSCANAQVYS 489
QY 490 LLPHPDPLPGSLVSIINADNGQLFALRSIDFELQAEFFRVGAAGSPALSSQALVRY 549
Db 490 LLPHPDPLPLASVSIINADNGHLFALRSIDYELQAEFFRVGATBPGSPALSSREALVRY 549
QY 550 LVADANDNAPFVLYPLQNGSAPCTEIVPRAAEAGYLVAKVYAVDGGSGONAMLSYOLLA 609
Db 550 LVADANDNSPFVLYPLQNGSAPCTEIVPRAAEPGYLVTKVYAVDGGSGONAMLSYOLLA 609
QY 610 TERGLFGVMAHNGEYRTA 627
Db 610 TERGLFGVMAHNGEYRTA 627

```

```

RESULT 12

```



QY 550 LVADANDNAPFVLYPLONGSAPCTELVPRAEAGYLVAKVAVDGSQGNAMLSYQLLKA 609  
Db 550 LVADANDNSPFVLYPLONGSAPCTELVPRAEAGYLVAKVAVDGSQGNAMLSYQLLKA 609  
QY 610 TERPGLFGWAHNGEVRTA 627  
Db 610 TERPGLFGWAHNGEVRTA 627

## RESULT 14

US-10-180-557-478  
Sequence 478, Application US/10180557  
Publication No. US20030022301A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C147  
CURRENT APPLICATION NUMBER: US/10/180,557  
PRIORITY FILING DATE: 2002-06-25  
Prior Application removed - See file Wrapper or Palm  
SEQ ID NO 478  
LENGTH: 800  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-180-557-478

Query Match 70.1%; Score 2249; DB 9; Length 800;  
Best Local Similarity 70.4%; Pred. No. 1, 2e-157;  
Matches 433; Conservative 70; Mismatches 113; Indels 0; Gaps 0;

QY 10 QKROVSLCYFLGVSWAGAPLRYFAEETERGFTLANLIDGLGEEELASAGCIYSD 69  
Db 10 RORVLEFLFVWGSLSLGSFGKRYSTEEKGSFVNLAKDKGLAEGLAAGTRVSD 69  
QY 70 ETIGELLNPLTGLDLNKLREELCGPTEPCVLPQDLLEKPFQIFRAELWVRDINDH 129  
Db 70 DNKQYLLDSHTGNLTLNEKLDREKLCGPKPCMLVFQILMDPFOIYRAELVRDINDH 129  
QY 130 SPVFLDEITLNLIESTPGATFLLESADSDVGINNLRNYTSSNVYPHINVDNGEGN 189  
Db 130 APVFQDETVLKISENTAEFTAERLERAODPDGLNGIQWYITSPNSFPHINISGDEGM 189  
QY 190 VSELVADKVLDEEVEVELRLTLGLDGSPPRSGLTLINLVDLNDNPEVESELYAV 249  
Db 190 IYELVADKVLDEEVEVELRLTLGLDGSPPRSGLTLINLVDLNDNPEVESELYAV 249  
QY 250 QVBNSEVSVLVTVSARDLDTGSNBEIYVAFYATERTLKTFRINSGNLTKAELNY 309  
Db 250 QVBNSEVSVLVTVSARDLDTGSNBEIYVAFYATERTLKTFRINSGNLTKAELNY 309  
QY 310 ELAIOYTLTQAKDGLSGKCTVYVHYVDINNPPELLMSLTSPENSPETVAVVR 369  
Db 310 ELVNSYINIQAMDGGSLSRCHVLEVDLNDNPELVSSFSNSVASENPETPLAVFK 369  
QY 370 IRDRSGNNAKVCSIODHLPFLVAKPSVENFYTLVTERALDREPEREYNTITVVDLGR 429  
Db 370 INDRSGNNAKVCSIODHLPFLVAKPSVENFYTLVTERALDREPEREYNTITVVDLGR 429  
QY 430 RLKTHNLVTVSDVNDNAPFESQTYTLTVRENNSPALHIGSVASTDDSGANQVYS 489  
Db 430 RLKTHNLVTVSDVNDNAPFESQTYTLTVRENNSPALHIGSVASTDDSGANQVYS 489

QY 490 LPPHPDPLGSLVSLNDNGOLFALRSIDFEALQAEFFRYGADRGSALSSQALVRY 549  
Db 490 LPPHPDPLGSLVSLNDNGOLFALRSIDFEALQAEFFRYGADRGSALSSQALVRY 549  
QY 550 LVADANDNAPFVLYPLONGSAPCTELVPRAEAGYLVAKVAVDGSQGNAMLSYQLLKA 609  
Db 550 LVADANDNSPFVLYPLONGSAPCTELVPRAEAGYLVAKVAVDGSQGNAMLSYQLLKA 609  
QY 610 TERPGLFGWAHNGEVRTA 627  
Db 610 TERPGLFGWAHNGEVRTA 627

## RESULT 15

US-09-931-836-52  
Sequence 52, Application US/09931836  
Publication No. US2003002749A1  
GENERAL INFORMATION:  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3030R1C1  
CURRENT APPLICATION NUMBER: US/09/931,836  
PRIORITY FILING DATE: 2001-08-16  
Prior Application removed - See file Wrapper or Palm  
SEQ ID NO 478  
LENGTH: 800  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-09-931-836-52

```

PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/132379
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/13383
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/135750
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/138166
PRIOR FILING DATE: 1999-06-08
PRIOR APPLICATION NUMBER: 60/144791
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/146970
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 09/311832
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/380142
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/644848
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 09/747259
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 09/816744
PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 09/854208
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/854280
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/874503
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: 09/869599
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/908,827
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: PCT/US99/10733
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30720
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: PCT/US00/05601
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/23522
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: PCT/US00/34956
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: 2001-08-26
PRIOR APPLICATION NUMBER: PCT/US01/17800
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: PCT/US01/19692
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: PCT/US01/21066
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: PCT/US01/21735
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 80
SEQ ID NO 52
LENGTH: 800
TYPE: PRT
ORGANISM: Homo Sapien
US-09-931-836-52

```

```

Query Match          70.1%; Score 2249; DB 9; Length 800;
Best Local Similarity 70.4%; Pred. No. 1 2e-157;
Matches 433; Conservative 70; Mismatches 113; Indels 0; Gaps 0;

QY 10 QKROVSLICVFLVSNAGAPRLRYFAETTERGFTLANLAIDGLVEELSGARCRIVSD 69
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 10 RQGVLFLEFWGVSLAGSGFRYSVTEETKEGSEYVNAKLKDGLAEGLAAGTRVIVSD 69
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 70 ETIFLLINPLTGDLNLNEKIDREELCGPPEPCVLPQQLLEKPFQFRELWVRINDH 129
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 70 DNKOYLLDSTGTNLNNEKIDREELCGPPEPCVLPQQLLEKPFQFRELWVRINDH 129
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 130 SPVFLDEKITLNIESTTGCATFLLESABSDVGINLRYNTSSNRYFHIYHNDGEGN 189
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 130 APVQDKEVYLIKISENTEAGFAFLEAQAOPDGLNIOYTTISPNFHIINISGDEGM 189
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 190 VYSELVLDKVLDRREVEBELRLTLTGDDGSPRSGTTLRIVLVDINDVPEVESLYKV 249
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 190 IYPELVLDKALDRREGELSLTLTALDGSRSRGTSTVRIVVDVNDNAPQFQALYET 249
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 250 QVPNSPVGSLVYVVSARDLDTGSGNELIYAAEFATERITLKTIRINSTGMLHKLKELNY 305
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 250 QAPENSPIGFLIVKVMEDVDGVAEVSSEFASINIRTTQINPFSGEITFLRELDY 309
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 310 EAIOTYTLTIOAKDGGISGKTVVHVHTDINDNPELMSLTSPIPENSPEYVAVFR 369
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 310 ELVNSYKINIQAMGGGGSARCRVLYEVDINDNPELIVSSFSNSVAENSPEYPLAVFK 369
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 370 IRRDGSNNAMKCSIODHPEVLKPSVENEFYLVTERALDREREITYNITTYVDLGT 429
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 370 INDRDSENGKMCVYIOENLEFLKPSVENEYIITEGALDRERAYENITTYVDLGT 429
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 430 RLKQHNLTIVVSVNNAPAFESOTYTLRYENNSPALHIGVSATDRDGSNAOYTVS 489
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 430 RLKTEHNTIVVSDVNDNAPAFETOTSTLEVRNNSPALHIGVSATDRDGSNAOYTVS 489
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 490 LLPHPDPLPGSLVSIINADNGQLFALRSLDFEALQAEFEVGAADGSPALSSQALVRY 549
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 490 LLPQDPLPLASLVSINADNGHFLALRSLDYELALQAEFEVVGATDGPALSSREALVRY 549
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 550 LVADANDNAPFVLYPLONGSAPCTELVPRAAEAGIYVAKVAVDGGSONAMLSYQLKA 609
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 550 LVLDANDNSPVLVPLONGSAPCTELVPRAAEAGIYVAKVAVDGGSONAMLSYQLKA 609
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 610 TEPGLFGVMAHNGEVRTA 627
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 610 TEPGLFGVMAHNGEVRTA 627
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: May 16, 2003, 10:48:51  
 Job time : 27 secs



GenCore version 5.1.4.P5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 16, 2003, 10:45:57 ; Search time 14 Seconds  
(Without alignments)  
1857.547 Million cell updates/sec

Title: US-10-050-704-125

Perfect score: 3208

Sequence: 1 MEARVHALKRGVSLICVF.....KATPGLGVANNGEVRTA 627

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2752	85.8	793	1	CDH7_HUMAN
2	2323	72.4	798	1	CDH2_HUMAN
3	2306	71.9	798	1	CDH6_HUMAN
4	2287	71.3	787	1	CDH5_HUMAN
5	2262	70.5	776	1	CDH3_HUMAN
6	2250	70.1	776	1	CDH8_HUMAN
7	2249	70.1	797	1	CDH9_HUMAN
8	2247	70.0	797	1	CDH4_HUMAN
9	2205	68.7	795	1	CDH1_HUMAN
10	2181	68.0	795	1	CDH11_HUMAN
11	2179.5	67.9	795	1	CDH10_HUMAN
12	2174.5	67.8	794	1	CDH12_HUMAN
13	2171	67.7	797	1	CDH13_HUMAN
14	2156.5	67.2	801	1	CDH14_HUMAN
15	2142	66.8	798	1	CDH15_HUMAN
16	2099.5	62.6	797	1	CDH16_HUMAN
17	1749.5	54.5	818	1	CDH17_HUMAN
18	1264.5	39.4	950	1	CDH18_HUMAN
19	1239.5	38.6	936	1	CDH19_HUMAN
20	1234.5	38.5	1007	1	CDH20_HUMAN
21	1230.5	38.4	950	1	CDH21_HUMAN
22	1227	38.2	950	1	CDH22_HUMAN
23	1222	38.1	950	1	CDH23_HUMAN
24	1203	37.5	937	1	CDH24_HUMAN
25	1200.5	37.4	947	1	CDH25_HUMAN
26	1200	37.4	950	1	CDH26_HUMAN
27	1194	37.2	950	1	CDH27_HUMAN
28	1192.5	37.0	949	1	CDH28_HUMAN
29	1188	36.8	963	1	CDH29_HUMAN
30	1179	36.8	941	1	CDH30_HUMAN
31	1175	36.6	950	1	CDH31_HUMAN
32	1157	36.1	948	1	CDH32_HUMAN
33	900.5	28.1	1180	1	PC12_MOUSE

34	880.5	27.4	1184	1	PC12_HUMAN
35	807	25.2	1069	1	PCD7_HUMAN
36	703	21.9	3298	1	PC16_HUMAN
37	603	18.8	3313	1	CLR3_RAT
38	599	18.7	3301	1	CLR3_MOUSE
39	598.5	18.7	3312	1	CLR3_HUMAN
40	597	18.6	5147	1	FAT_DROME
41	595.5	18.6	2923	1	CLR2_HUMAN
42	590	18.4	3014	1	CLR1_HUMAN
43	589	18.4	3579	1	STAN_DROME
44	585	18.2	3354	1	CADN_HUMAN
45	584	18.2	3034	1	CLR1_MOUSE

## ALIGNMENTS

RESULT 1	ID	CDH7_HUMAN	STANDARD:	PRT:	793 AA.
AC	Q9Y5E2	CDH7_HUMAN	STANDARD:	PRT:	793 AA.
DT	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Protocadherin beta 7 precursor (PCDH-beta7).				
GN	PCDH7.				
OS	Homo sapiens (human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RP	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99308636; PubMed=10380929;				
RA	Wu Q., Mandatis T.;				
RT	"A striking organization of a large family of human neural cadherin-				
RT	like cell adhesion genes.";				
RL	Cell 97:779-790(1999).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=21223055; PubMed=11322959;				
RA	Vanhaest K., Koels P., Vanden Eynde E., van Roy F.;				
RT	"The human and murine protocadherin-beta one-exon gene families show				
RT	high evolutionary conservation, despite the difference in gene				
RT	number.";				
RL	FEBS Lett. 495:120-125(2001).				
CC	- FUNCTION: POTENTIAL CALCIUM-DEPENDENT CELL-ADHESION PROTEIN.				
CC	- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).				
CC	- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).				
CC	-----				
DR	EMBL: AF152500; AAD43761.1; -				
DR	EMBL: AF217750; AAK51617.1; -				
DR	HSSP: P15116; INCU.				
DR	GeneW: HGNC:8692; PCDH7.				
DR	MIM: 606333; -				
DR	MIM: 604967; -				
DR	InterPro: IPR002126; Cadherin.				
DR	Pfam: PF00028; cadherin.5.				
DR	PRINTS: PRO0205; CADHERIN.				
DR	SMART: SM00112; CA: 5.				
DR	PROSITE: PS00332; CADHERIN_1; 5.				
DR	PROSITE: PS00368; CADHERIN_2; 5.				
KW	Calcium-binding; Cell adhesion; Glycoprotein; Signal; Repeat;				
KW	Transmembrane; Multigene family.				
FT	SIGNAL	1	26	POTENTIAL.	
FT	CHAIN	27	793	PROTOCOLADHERIN BETA 7.	

FT DOMAIN 27 688 EXTRACELLULAR (POTENTIAL).  
 FT TRASMEN 689 709 POTENTIAL.  
 FT DOMAIN 710 793 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 35 133 CADHERIN 1.  
 FT DOMAIN 138 242 CADHERIN 2.  
 FT DOMAIN 247 347 CADHERIN 3.  
 FT DOMAIN 352 451 CADHERIN 4.  
 FT DOMAIN 456 561 CADHERIN 5.  
 FT DOMAIN 568 671 CADHERIN 6.  
 FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 436 436 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 793 AA: 86706 MW: 856770986530C755 CRC64;

Query Match 85.8%; Score 2752; DB 1; Length 793;  
 Best Local Similarity 85.3%; Pred. No. 3,9e-16;  
 Matches 535; Conservative 34; Mismatches 58; Indels 0; Gaps 0;

QY 1 MEARVVALOKROVSLICVFLGVSWAGAPLRYFAETEGTFLANLADLGVEELS 60  
 DB 1 MEARVVALOKROVSLICVFLGVSWAGAPLRYFAETEGTFLANLADLGVEELS 60  
 QY 61 ARGRIVSDITIGFLNPLTGLDLNKLREELCGTPECVLPFOLLKPKPQIFRAE 120  
 DB 61 ARGRIVSDITIGFLNPLTGLDLNKLREELCGTPECVLPFOLLKPKPQIFRAE 120  
 QY 121 LMRVINDHSPYFLDREITLNLSTTPGATFLLESADSDVGNLNNYTISSNVYFHI 180  
 DB 121 LMRVINDHSPYFLDREITLNLSTTPGATFLLESADSDVGNLNNYTISSNVYFHI 180  
 QY 181 NVHDGEGNVSELYLDVLDREYVELRLTLTGDDGSPRSGLTIRIIVLINDNVP 240  
 DB 181 NVHDGEGNVSELYLDVLDREYVELRLTLTGDDGSPRSGLTIRIIVLINDNVP 240  
 QY 241 EFVESLYKYOVSPENSVGLVTVSARDLDTGNGEIVYAFYATERTLKTFRINSTGN 300  
 DB 241 EFVESLYKYOVSPENSVGLVTVSARDLDTGNGEIVYAFYATERTLKTFRINSTGN 300  
 QY 301 LHLKELNVEALQYTLTITQAKDGGSLGKCTVVVHTVDINDNPELIMSLTSPIDENS 360  
 DB 301 LHLKELNVEALQYTLTITQAKDGGSLGKCTVVVHTVDINDNPELIMSLTSPIDENS 360  
 QY 361 PETVAVERIRIDRSGNNAKWCSTQDHLPEVLKPSVNEFTYTERALDDEERTENIT 420  
 DB 361 PETVAVERIRIDRSGNNAKWCSTQDHLPEVLKPSVNEFTYTERALDDEERTENIT 420  
 QY 421 ITVVDLGPRLKTOHNLTVTVSDVNDNAPTESQTYTLTVRENNSPALHIGSVSATDRDS 480  
 DB 421 ITVVDLGPRLKTOHNLTVTVSDVNDNAPTESQTYTLTVRENNSPALHIGSVSATDRDS 480  
 QY 481 GANOVWYSLRPHDPOPLGSIYSINDNQLFALRSLDEALQAEFRGADRSSPA 540  
 DB 481 GANOVWYSLRPHDPOPLGSIYSINDNQLFALRSLDEALQAEFRGADRSSPA 540  
 QY 541 LSSQALVAVLADANDNAPFVLYPLONGSAPCTELVPAAGLVLKVAVVDSDSONA 600  
 DB 541 LSSQALVAVLADANDNAPFVLYPLONGSAPCTELVPAAGLVLKVAVVDSDSONA 600  
 QY 601 WLSTYLLKATEPGLFGVAHNGEVRTA 627  
 DB 601 WLSTYLLKATEPGLFGVAHNGEVRTA 627

RESULT 2

ID CDB2\_HUMAN STANDARD; PRT; 798 AA.  
 AC Q9Y5E7;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Protocadherin beta 2 precursor (PCDH-beta2).  
 GN PCDHB2.

OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99308636; PubMed=10380929;  
 RA Wu Q., Maniatis T.;  
 RT "A striking organization of a large family of human neural cadherin-  
 RL like cell adhesion genes.";  
 RL Cell 97:779-790(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2123035; PubMed=11322959;  
 RA Vanhalst K., Kools P., Vanden Eynde E., van Roy F.;  
 RT "The human and murine protocadherin-beta one-exon gene families show  
 RT high evolutionary conservation, despite the difference in gene  
 RT number.";  
 RL PNAS 98:495-499(2001).  
 RL -1- FUNCTION: POTENTIAL CALCIUM-DEPENDENT CELL-ADHESION PROTEIN.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: AF152495; AAD43756.1; -;  
 DR EMBL: AF217756; AAK51623.1; -;  
 DR HSSP: P15116; INCU.  
 DR GeneW: HGNC:8687; PCDHB2.  
 DR MIM: 606328; -;  
 DR MIM: 604967; -;  
 DR InterPro: IPR002126; Cadherin.  
 DR Pfam: PF00028; cadherin; 5.  
 DR PRINTS: PRO0205; CADHERIN.  
 DR SMART: SM00112; CA; 6.  
 DR PROSITE: PS00232; CADHERIN\_1; 5.  
 DR PROSITE: PS00268; CADHERIN\_2; 6.  
 KW Calcium-binding; Cell adhesion; Glycoprotein; Signal; Repeat;  
 KW Transmembrane; Multigene family.  
 FT SIGNAL 1 30  
 FT CHAIN 31 798  
 FT DOMAIN 31 692  
 FT TRANSMEM 693 713  
 FT DOMAIN 714 798  
 FT DOMAIN 136 244  
 FT DOMAIN 249 349  
 FT DOMAIN 354 453  
 FT DOMAIN 458 563  
 FT DOMAIN 570 673  
 FT CARBOHYD 171 171  
 FT CARBOHYD 420 420  
 FT CARBOHYD 438 438  
 FT CARBOHYD 569 569  
 SQ SEQUENCE 798 AA: 87253 MW: 82602865378CB27 CRC64;

Query Match 72.4%; Score 2323; DB 1; Length 798;  
 Best Local Similarity 73.0%; Pred. No. 1.7e-14;  
 Matches 451; Conservative 65; Mismatches 102; Indels 0; Gaps 0;

QY 10 OKROVSLICVFLGVSWAGAPLRYFAETEGTFLANLADLGVEELSARGRIYSD 69  
 DB 10 OKROVSLICVFLGVSWAGAPLRYFAETEGTFLANLADLGVEELSARGRIYSD 69  
 QY 12 KOROVILFVVLGIGASQOPRHYSVAETEGSVANLKLGLGIELAVRGARYSK 71  
 DB 12 KOROVILFVVLGIGASQOPRHYSVAETEGSVANLKLGLGIELAVRGARYSK 71  
 QY 70 EITGFLNPLTGLDLNKLREELCGTPECVLPFOLLKPKPQIFRAELIWDINDH 129  
 DB 70 EITGFLNPLTGLDLNKLREELCGTPECVLPFOLLKPKPQIFRAELIWDINDH 129  
 DB 72 GKMHILQFDROGDLNKLREELCGTPECVLPFOLLKPKPQIFRAELIWDINDH 131



```

QY 130 SPVFLDREITLNIESTTPGATFLLESASHDSVGINNLRNTYSSNVYFHIWVHNGEEN 189
DB 132 SPVFLKEILLKIPESITPTCTFLIERADLDVGTNSLQNTYISPFHFMNQSDLDGI 191
QY 190 VSEVLVDKLVDEEVEPELRLTLTGIDGSPPSGTTLRILVLDINDVPEVESLYK 249
DB 192 ILPOLVLRNLRDREPELRLTLTGIDGSPPSGTTLRILVLDINDVPEVESLYK 251
QY 250 QVPEPVSGLVYVSARDLDGSGEIVYAFYATERLTKPRINSTGNLHKELNY 309
DB 252 QIPEDSPVSGOVAIVSARDLDIGTNEISYASQASEDIRKTFRLSAKSGELLROKIDF 311
QY 310 EAIQVYTLTIQAKDGGGLSGCTVYVYVYDINDNPELMSLTSPINSEPEVAVPR 369
DB 312 ESIGTYVNIQATDGGGLSGCTVYVYVYDINDNPELMSLTSPINSEPEVAVPR 371
QY 370 IRRDGSNNAKMWCSTIODHLPEVLKSVENFTYTLVTERALDREERTYITVTDLGP 429
DB 372 VSDPDGSGNRMVCSIODDLPEFLKPSVENFTYTLVTERALDREERTYITVTDLGP 431
QY 430 RLKTOHNLVTVSDVNDNAPTESQTYTLVLRVRENNSPALHIGSVSATDSDGAMQVYS 489
DB 432 RLKTEINITYLVSDVNDNAPTESQTYTLVLRVRENNSPALHIGSVSATDSDGAMQVYS 491
QY 490 LRPHPDQLPLGSLVSLNADNGQLFALRLDPEALQAFERVGAADRGSPALSSQALVY 549
DB 492 LRPDPHPPLASVSLNADNGQLFALRLDPEALQAFERVGAADRGSPALSSQALVY 551
QY 550 LVADANDNAPFVLYPLONSAPCTELVPRAEPGYLVAVVAVDGSNGAMLSIOLKA 609
DB 552 LVADANDNAPFVLYPLONSAPCTELVPRAEPGYLVAVVAVDGSNGAMLSIOLKA 611
QY 610 TERPLFGVMAHNGEVRTA 627
DB 612 TERPLFGVMAHNGEVRTA 629

```

RESULT 3  
CDBE\_HUMAN STANDARD: PRI: 798 AA.

AC 09Y5E9;  
DT 16-OCT-2001 (Rel. 40. Created)  
DT 15-OCT-2001 (Rel. 40. Last sequence update)  
DT 15-JUN-2002 (Rel. 41. Last annotation update)  
DE Protocadherin beta 14 precursor (PCDH-beta14).  
GN PCDHB14.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9308636; PubMed=10380929;  
RA Wu O., Maniatis T.;  
RT "A striking organization of a large family of human neural cadherin-  
RT like cell adhesion genes.";  
RT Cell 97:779-790(1999).  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=2123055; PubMed=11322959;  
RA Vanhaesbroeck B., Koois P., Vanden Eynde E., van Roy F.;  
RT "The human and murine protocadherin-beta one-exon gene families show  
RT high evolutionary conservation, despite the difference in gene  
RT number.";  
RT FEBS Lett. 495:120-125(2001).  
CC -1- FUNCTION: POTENTIAL CALCIUM-DEPENDENT CELL-ADHESION PROTEIN.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (by similarity).  
CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: AF152493; AAD43754.1;  
CC EMBL: AF217744; AAK51612.1;  
CC HSSP: P51116; INCU.  
CC Genes: HGNC:6685; PCDHB14.  
CC MIM: 606340; -;  
DR MIM: 604967; -;  
DR InterPro: IPR002126; Cadherin.  
DR Pfam: PF00028; cadherin; 5.  
DR PRINTS: PR00205; CADHERIN.  
DR SMART: SM00112; CA: 5.  
DR PROSITE: PS00232; CADHERIN\_1; 5.  
DR PROSITE: PS50268; CADHERIN\_2; 5.  
KW Calcium-binding; Cell adhesion; Glycoprotein; Signal; Repeat;  
KW Transmembrane; Multigene family.  
FT SIGNAL 1 26  
FT CHAIN 27 798  
FT DOMAIN 27 686  
FT TRANSMEM 687 711  
FT DOMAIN 712 798  
FT DOMAIN 35 133  
FT DOMAIN 138 242  
FT DOMAIN 247 347  
FT DOMAIN 352 451  
FT DOMAIN 456 561  
FT DOMAIN 568 671  
FT CARBOHYD 169 169  
FT CARBOHYD 359 359  
FT CARBOHYD 418 418  
FT CARBOHYD 436 436  
FT CARBOHYD 487 487  
FT CARBOHYD 567 567  
SQ SEQUENCE 798 AA; 87547 MW; 65151C1A423BB52E CRC64;

Query Match 71.9%; Score 2306; DB 1; Length 798;  
Best Local Similarity 71.3%; Pred. No. 2.3e-146;  
Matches 447; Conservative 78; Mismatches 102; Indels 0; Gaps 0;

QY 1 MEARYVHALQROYSLCLVFLGVSWAGAPLRYFAETERTKGTFLANLADLGIVEELS 60  
DB 1 MEIRALDLRRKROYLIFVLVGLISRAGTESAHYSAEBTEIGSVANLARDLGIVEELS 60  
QY 61 ARGGRVSDENITGELLNPLTGDLINKLDRBELCGTEFCVPLQOLLEKPPQITRAE 120  
DB 61 SREARVVSDDNKRYLHLDLTLGNLLINKLDRDELCSGTEPCVILHFOVLENPLOFFRFE 120  
QY 121 LVRVDINDSHSPFLDREITLNIESTTPGATFLLESASHDSVGINNLRNTYSSNVYFHI 180  
DB 121 LCVKDINDSHSPFLDREITLNIESTTPGATFLLESASHDSVGINNLRNTYSSNVYFHI 180  
QY 181 NVHDNGBGVNYSVLVDKLVDEEVEPELRLTLTGIDGSPPSGTTLRILVLDINDV 240  
DB 181 KIPDSSDKRIVPELVLDRLDYEQAELRLTLTVADGSPSPKSGTTLVLRKLDINDNAP 240  
QY 241 EPEVSLYKVOYPEPVSGLVYVSARDLDGSGEIVYAFYATERLTKPRINSTGN 300  
DB 241 EPPSLYEVQYPEPVSGLVYVSARDLDGSGEIVYAFYATERLTKPRINSTGN 300  
QY 301 LHLKAEVLEAIQTYTLTIQAKDGGGLSGCTVYVYVYDINDNPELMSLTSPINSE 360  
DB 301 VNLSPDLDFEYIQSYTNIQATDGGGLSGCTVYVYVYDINDNPELMSLTSPINSE 360  
QY 361 PEYVAVAFRIIDRDSGNNAKWCSTIODHLPEVLKSVENFTYTLVTERALDREERTY 420  
DB 361 SETVLAFLSIIDDSGNGRMICSTIODHLPEFLKPSVENFTYTLVTERALDREERTY 420  
QY 421 ITVVDLGPRLKTEYNITVTLSDVNDNAPTESQTYTLVLRVRENNSPALHIGSVSA 480  
DB 421 ITVVDLGPRLKTEYNITVTLSDVNDNAPTESQTYTLVLRVRENNSPALHIGSVSA 480

QY 481 GANAQVTSLLPDPHPLGLSIVSINADNGOLFALRSIDLEALQAEFFRVGAADRGSPA 540  
 DB 481 GINAQVNSILPDPDRHPLPLASIVSINADNGHFLRLSLDLEALQAEFFRVGAADRGSPA 540  
 QY 541 ISSQALVRYLVADANDANPFLVYLPYQNGSAPCTELVPRAPAGYLVAKVAVDGDGQNA 600  
 DB 541 LSSEALVRYLVADANDNSPFVYLPYQNGSAPCTELVPRAPAGYLVAKVAVDGDGQNA 600  
 QY 601 WLSTYQLKATPEGLFGVMAHNGEVRTA 627  
 DB 601 WLSTYQLKATPEGLFGVMAHNGEVRTA 627

RESULT 4  
 CDBF\_HUMAN STANDARD: PRT: 787 AA.  
 AC Q9Y5E6: 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Protocadherin beta 15 precursor (PCDH-beta15).  
 GN PCDHB15.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99308636; PubMed=10380929;  
 RA Wu Q., Mantatis T.;  
 RT "A striking organization of a large family of human neural cadherin-  
 RT like cell adhesion genes";  
 RL Cell 97:779-790(1999).  
 RP [2]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21223055; PubMed=11322959;  
 RA Vanhalst K., Kools P., Vanden Eynde E., van Roy F.;  
 RT "The human and murine protocadherin-beta one-exon gene families show  
 RT high evolutionary conservation, despite the difference in gene  
 RT number";  
 RL FEBS Lett. 495:120-125(2001).  
 CC -1- FUNCTION: POTENTIAL CALCIUM-DEPENDENT CELL-ADHESION PROTEIN.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and thus statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: AF152494; AAD43755.1; -  
 DR EMBL: AF217742; AAK51610.1; -  
 DR Genew: HGNC:8686; PCDHB15.  
 DR MIM: 606341; -  
 DR MIM: 604967; -  
 DR InterPro: IPR002126; Cadherin.  
 DR Pfam: PF00028; cadherin.5.  
 DR PRINTS: PR00205; CADHERIN.  
 DR SMART: SM00112; CA.5.  
 DR PROSITE: PS00232; CADHERIN.1; 5.  
 DR PROSITE: PS50268; CADHERIN.2; 6.  
 KW Calcium-binding; Cell adhesion; Glycoprotein; Signal; Repeat;  
 KW Transmembrane; Multigene family.  
 FT SIGNAL 1 26 POTENTIAL.  
 FT CHAIN 27 787 PROTOCADHERIN BETA 15.  
 FT DOMAIN 27 690 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 691 711 POTENTIAL.  
 FT DOMAIN 712 787 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 133 CADHERIN 1.

FT DOMAIN 138 242 CADHERIN 2.  
 FT DOMAIN 247 347 CADHERIN 3.  
 FT DOMAIN 352 451 CADHERIN 4.  
 FT DOMAIN 456 561 CADHERIN 5.  
 FT DOMAIN 568 671 CADHERIN 6.  
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 787 AA; 86328 MW; 8DB8D3A07B0BA56 CRC64.

Query Match 71.3%; Score 2287; DB 1; Length 787;  
 Best Local Similarity 72.3%; Pred. No. 4, 2e-145; Indels 0; Gaps 0;  
 Matches 447; Conservative 67; Mismatches 104; Indels 0; Gaps 0;

QY 10 QKROVSLICVFLGVSNAGAEPRLRYFAEETERTGLFANLAIDGLGVEELASRGRIYSD 69  
 DB 10 EQQVILILLLEVLTLAGMEPRRYSMETFERGSEFVANLANLDGLGVGLARGARVYSE 69  
 QY 70 ETIGFLINPLTDDLLNEKLDREELCGTPREVLPQLLKEKPPQIPFAELVAVDINDH 129  
 DB 70 DNEQGLQDLQTLINLEKLDREELCGTPREIMHFVLLKPLEVFRAELLVTDINDH 129  
 QY 130 SPVFLDREITLNTLESTTFCATPLESADSDVGINNLNNTYISSNVYHINVHNGEGN 189  
 DB 130 SPEFEREMTKLPETSSISLGTVPPLKARDLVGSMNQNINSPSHHVSTRRGDSR 189  
 QY 190 VSEFLVDKYLDEEVEPELRLTLTGLDGSPRSGTTLIRILVLDINDVPEVESLYKV 249  
 DB 190 KPELVLDPELDREEOAEELRLTLTAVDGGSPRSGTQVLLILVLDANDNAPEFVQALYEV 249  
 QY 250 QVPENSPVGLVTVYVAROLDGSGEITYAFYATERTLKTFRINSTGNHLKAEINX 309  
 DB 250 QVPENSPVGLVTVYVAROLDGSGEITYAFYATERTLKTFRINSTGNHLKAEINX 309  
 QY 310 EAIQVTLTFOAKDGGGLSGKCTVAVVVDINDNPBELMSSITSPENSPETVAVER 369  
 DB 310 EFMSSVDLDEASDGGGLSGKCSVSKVLDVNDNPELISITLSPINSETEVALFR 369  
 QY 370 IRDRSGNNAKWCSIQDHLPEVLKPSVENEFTLVTERALDRBERTEVNTITVLDLGP 429  
 DB 370 IRDRSGNNAKWCSIQDHLPEVLKPSVENEFTLVTERALDRBERTEVNTITVLDLGP 429  
 QY 430 RLKTHNLTAVTVSDVNDNAPTFSQTYTLRVPRNNSPALHISVSATDSDGANNQVYS 489  
 DB 430 RLKTHNLTAVTVSDVNDNAPTFSQTYTLRVPRNNSPALHISVSATDSDGANNQVYS 489  
 QY 490 LRPDPHPLGLSIVSINADNGOLFALRSIDLEALQAEFFRVGAADRGSPALSVAV 549  
 DB 490 LRPDPHPLGLSIVSINADNGOLFALRSIDLEALQAEFFRVGAADRGSPALSVAV 549  
 QY 550 LVADANDNAPFLVYLPYQNGSAPCTELVPRAPAGYLVAVAVDGDGQNAVLSTQLKA 609  
 DB 550 LVADANDNAPFLVYLPYQNGSAPCTELVPRAPAGYLVAVAVDGDGQNAVLSTQLKA 609  
 QY 610 TERGLFGVMAHNGEVRTA 627  
 DB 610 TERGLFGVMAHNGEVRTA 627

RESULT 5  
 CDB3\_HUMAN STANDARD: PRT: 796 AA.  
 AC Q9Y5E6: 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Protocadherin beta 3 precursor (PCDH-beta3).  
 GN PCDHB3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.

Db 181 LTRSRDRKRKPELVLDALDREQPELSLTFLTALDGGSPSSGTAQINIQVLDINDNP 240

Qy 241 EFVESLTVYQVPEPNSPVASIVYTSVARLDPGSGNCEIYAFATERTLKTFRINSTGN 300

Db 241 EFAQPLIEVAALIENTPVASIVYTSASDLDGSGFETISTIAFPHASEETRKTFQLNPTGD 300

Qy 301 LHLKALNVEALIQYVTLTLTQAKDGGGSGKCTVVVHVTDINDNPPELLMSLTSPSPENS 360

Db 301 MQLVYLNFEEALINSYEVADIEAKDGGGSGKSTVIYQVVDVNDNPPELLTSSVNSPIENS 360

Qy 361 PETVAVARIRIDROSGNNAKAWVCISQDHLPEFLAKSVENFTYLVTERALDREEREYNT 420

Db 361 GETVAVAVSVSDSGDNGRWCSIEENLPFLKPSVENFTYLVSEGALDRETRREYNT 420

Qy 421 ITVVDLGTPLRKTQHNLTIVTSDVDNAPTEFSQTYTTLVRFNNSPALHIGSVATDRDS 480

Db 421 ITITDLGPRLTKTKNITIVLSDVDNAPAFQIISTYTLFVRNNSPALHIGSVATDRDS 480

Qy 481 CANAOVYTSLLPRHPDPLGSLVSINADNGOLFALRSIDFEALQAFEPFVGAADRGSPA 540

Db 481 GTNAOVYTSLLPRHPDPLGSLVSINADNGOLFALRSIDFEALQAFEPFVGAADRGSPA 540

Qy 541 LSSQALVAVLVADADNDNPFVLYPLQNSAPCTELVPAAGEGYVAVVAVVADGSGONA 600

Db 541 LSSELAVVAVLVADANDNPFVLYPLQNSAPCTELVPAAGEGYVAVVAVVADGSGONA 600

Qy 601 WLSTYQLKATEPEGLFGVVAHNGEVETA 627

Db 601 WLSTYQLKATEPEGLFGVVAHNGEVETA 627

RESULT 6

CDBG\_HUMAN STANDARD; PRT; 776 AA.

AC Q9NRJ7; Q9HCF1; 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Protocadherin beta 16 precursor (PCDH-beta16) (Protocadherin 3x).

GN PCDHb16 OR PCDH3X OR KIAA1621.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxId=9606; [1]

RN [1] SEQUENCE FROM N.A.

RP Kools P.F.J., Van Roy F.; "The human protocadherin 3 gene cluster, expression analysis of these one-exon genes"; Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases. [2]

RN [2] SEQUENCE FROM N.A.

RP TISSUE=Brain; MEDLINE=20450683; PubMed=10997877; Nagase T., Kikuno R., Nakayama M., Hirotsawa M., Ohara O.; "Prediction of the coding sequences of unidentified human genes. XVIII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro."; DNA Res. 7:273-281(2000). [3]

RN [3] SEQUENCE FROM N.A.

RP Kools P.F.J., van Roy F.M.; "Molecular analysis of the human protocadherin-3 (PCDH-beta) gene cluster"; Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.

CC -!- FUNCTION. POTENTIAL CALCIUM-DEPENDENT CELL-ADHESION PROTEIN.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).

CC -!- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF217757; AAF81914.1; -  
 DR EMBL: AB046841; BAB13447.1; ALT\_INIT.  
 DR EMBL: AF282973; AAC10030.1; -  
 DR Genew: HGNC:14546; PCDBH16.  
 DR MIM: 606345; -  
 DR MIM: 604967; -  
 DR InterPro: IPR002126; Cadherin.  
 DR Pfam: PF00028; cadherin.5.  
 DR PRINTS: PR00205; CADHERIN.  
 DR SMART: SM00112; CA: 5.  
 DR PROSITE: PS00232; CADHERIN\_1; 5.  
 DR PROSITE: PS0268; CADHERIN\_2; 6.  
 DR KEGG: K04400; Cell adhesion; Glycoprotein; Signal; Repeat;  
 KW Transmembrane; Multigene family.  
 FT SIGNAL 1 28  
 FT CHAIN 1 28  
 FT DOMAIN 29 776  
 FT TRANSMEM 29 690  
 FT TRANSMEM 691 711  
 FT DOMAIN 712 776  
 FT DOMAIN 35 133  
 FT DOMAIN 138 242  
 FT DOMAIN 247 347  
 FT DOMAIN 352 451  
 FT DOMAIN 456 561  
 FT DOMAIN 568 671  
 FT CARBOHD 418 418  
 FT CARBOHD 436 436  
 FT CARBOHD 567 567  
 FT CONFLICT 482 482  
 FT CONFLICT 525 526  
 FT CONFLICT 532 532  
 FT CONFLICT 543 543  
 FT CONFLICT 638 638  
 FT CONFLICT 652 652  
 FT CONFLICT 710 710  
 FT SEQUENCE 776 AA; 84935 MW; 93EFC641ADE0FB CRC64;  
 Query Match 70.1%; Score 2250; DB 1; Length 776;  
 Best Local Similarity 69.5%; Pred. No. 1.2e-142;  
 Matches 436; Conservative 74; Mismatches 117; Indels 0; Gaps 0;

QY 1 MEARVHALQKROVSLICVPLGVSWAGAPLRYVAEETERTGFLANLADLGIVBELS 60  
 DB 1 MEIGMHRNRROVYLVFVLLSLISGAGALGSVYETERTGFLANLADLGIVBELS 60  
 QY 61 ARGCIVSDERTIGFLNPLNPLGDLINKLRELCGTEPCVLPOLLLEKPOITRAE 120  
 DB 61 TRKATISQGNKRLQLAQGDLLINKLRELCGTEPCVLPOLLLEKPOITRAE 120  
 QY 121 LWRDINDHSPVFLDREITLILSTPGATFLLESADSDVGINLNRYTSSNYFEH 180  
 DB 121 LRAVIDINHSPVFLDREITLILSTPGATFLLESADSDVGINLNRYTSSNYFEH 180  
 QY 121 LRAVIDINHSPVFLDREITLILSTPGATFLLESADSDVGINLNRYTSSNYFEH 180  
 DB 121 LRAVIDINHSPVFLDREITLILSTPGATFLLESADSDVGINLNRYTSSNYFEH 180  
 QY 161 NVHDNGEENYSELVLDVLDREYVPELRLTLTGDDGSPRSGLTIRILVINDNVP 240  
 DB 161 NVHDNGEENYSELVLDVLDREYVPELRLTLTGDDGSPRSGLTIRILVINDNVP 240  
 QY 181 LIHEFRDRKRYPELVLDELREPEPRLTLTALDGGSPRSGLTIRILVINDNVP 240  
 DB 181 LIHEFRDRKRYPELVLDELREPEPRLTLTALDGGSPRSGLTIRILVINDNVP 240  
 QY 241 FEVESLYVQVPEPNSVGLVTVTSARDLDSNGCEITVYAFYETERTKFRINSN 300  
 DB 241 FEVESLYVQVPEPNSVGLVTVTSARDLDSNGCEITVYAFYETERTKFRINSN 300  
 QY 301 LHLKALNVEAIIQYTLTQAKDGGSLGKCTVVVYVHDINDNPELMSLTPIPENS 360  
 DB 301 LHLKALNVEAIIQYTLTQAKDGGSLGKCTVVVYVHDINDNPELMSLTPIPENS 360  
 QY 301 VRLRKQVDFEWMVTSYEVRIKATDGGSLGKCTLLQVVDVNDPNQVYTMALTPIPENS 360  
 DB 301 VRLRKQVDFEWMVTSYEVRIKATDGGSLGKCTLLQVVDVNDPNQVYTMALTPIPENS 360  
 QY 361 PETVAVRIRINDRSGNNAKVCISIDHLPVLKPSYENFTVTERALDREERTENIT 420  
 DB 361 PETVAVRIRINDRSGNNAKVCISIDHLPVLKPSYENFTVTERALDREERTENIT 420  
 QY 361 PETVAVRIRINDRSGNNAKVCISIDHLPVLKPSYENFTVTERALDREERTENIT 420  
 DB 361 PETVAVRIRINDRSGNNAKVCISIDHLPVLKPSYENFTVTERALDREERTENIT 420

QY 421 ITVTDGTPRLKTOHNLTVYSDVNDNAPTFQSTYTLFRVNNSPALHIGSVATDRDS 480  
 DB 421 ITVTDGTPRLKTOHNLTVYSDVNDNAPTFQSTYTLFRVNNSPALHIGSVATDRDS 480  
 QY 481 GAAQVYTSILPPHDPQLPLGSLVSTNANGOLFALRSIDFALQAFERRVCAARGSPA 540  
 DB 481 GAAQVYTSILPPHDPQLPLGSLVSTNANGOLFALRSIDFALQAFERRVCAARGSPA 540  
 QY 541 LSSQALVRLVADANAPFVLYPLONGSAPCTELVPRAAGVYAVKVAVDGDSGNA 600  
 DB 541 LSSQALVRLVADANAPFVLYPLONGSAPCTELVPRAAGVYAVKVAVDGDSGNA 600  
 QY 601 WLSYQLKATEPGLFGVMAHNGEVRTA 627  
 DB 601 WLSYQLKATEPGLFGVMAHNGEVRTA 627

## RESULT 7

ID CDBA\_HUMAN STANDARD; PRT; 800 AA.  
 AC 09UN67;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Protecadherin beta 10 precursor (PCDH-beta10).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NCBI\_TaxID:9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE-99308636; PubMed-10380929;  
 RA Wu Q., Maniatis T.;  
 RT "A striking organization of a large family of human neural cadherin-  
 like cell adhesion genes."  
 RL Cell 97:779-790(1999).  
 CC -1- FUNCTION: POTENTIAL CALCIUM-DEPENDENT CELL-ADHESION PROTEIN.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (by similarity).  
 CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



```

Db 61 ARCTRYVSDNNKQYLLLDSTGKLLTNEKLDKRLKCGPEPCMLTYFQILMDPPFOIYRAE 120
QY 121 LWRVDINDSPVFLDREITLNLLESTPGATLLESABSDVGINLKNYITSSNYFEH 180
Db 121 LVRVDINDSPVFLRHEKEMWLKISENTAEGTAFLERAOPDEGSHNSIQWYITSSNFEH 180
QY 181 NVHNDGEGNSLTVLDKVLDRREVEPELRLTLTGDSGSPRSCTTLRLIVLDINDNP 240
Db 181 KIGSGDEGATYPELVLDKDLDRDEBELSLTLALDGGSPSSGSTTRIVYLDVNDAP 240
QY 241 EYVESLYKQVPEPNSPVGLVYVSARDLDTGSNGEIVAFYATERTLTPRINSTGN 300
Db 241 QFAQALYETQAPENSPVGLIYKVSAGDADSGVNAEVSSEFDASBDILTTQINPFGE 300
QY 301 LHKALNTEALQTYTLTQADGGGLSKCTVYVHVTDINDNPPELLMSSITSPENS 360
Db 301 IFRRELDYELVNSYKINQADGGGLSARCTVLKVLDSNNPELITSSLSNVAENS 360
QY 361 PETVAVFRIIRDSDGNNAKWCYSIODHLPFLKPSVENFYTLVTERALDREREYNT 420
Db 361 PGVILAVFKIKDRDSEKNTCTCYVDNLPEFLKPSVDNFYTLMEGALDRSKAEYNT 420
QY 421 IYVTDLGPRRLKQNLVTVSDVNDNAPFESQTTTYTLRVRENSPALHIGSVATDRS 480
Db 421 IYVTDLGPRRLKTEHSTLQVSDVNDNAPAFQTSTYTLFVRENSPALHIGSVATDRS 480
QY 481 GNAQVTVSLPRPHDPLGLSVISINDNGOLFALRSIDFALDAFERGADRGSPA 540
Db 481 GNAQVTVSLPRPHDPLGLSVISINDNGOLFALRSIDFALDAFERGADRGSPA 540
QY 541 LSSQALVRLVADANDNAPFVLYPLONGSAPCTELVPRAEAGYVAVVAVDGGGNA 600
Db 541 LSSQALVRLVADANDNAPFVLYPLONGSAPCTELVPRAEAGYVAVVAVDGGGNA 600
QY 601 WLSYOLKATKEGELGVMAHNGEVRTA 627
Db 601 WLSYOLKATKEGELGVMAHNGEVRTA 627

RESULT 9
CDB4_HUMAN STANDARD: PRT; 795 AA.
ID AC Q9Y5E5;
DR 16-OCT-2001 (Rel. 40, Created)
DR 16-OCT-2001 (Rel. 40, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protocadherin beta 4 precursor (PCDH-beta4).
GN PCDH4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99308636; PubMed=10380929;
RA Wu Q., Maniatis T.;
RT "A striking organization of a large family of human neural cadherin-
RT like cell adhesion genes.";
RL Cell 97:779-790(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21223055; PubMed=11322959;
RA Vanhaest K., Kools P., Vanden Eynde E., van Roy F.;
RT "The human and murine protocadherin-beta one-exon gene families show
RT high evolutionary conservation, despite the difference in gene
RT number.";
RL FEBS Lett. 495:120-125(2001).
CC -1- FUNCTION: POTENTIAL CALCIUM-DEPENDENT CELL-ADHESION PROTEIN.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (by similarity).
CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; AF152497; AAD43758.1; -
CC EMBL; AF217754; AAK51621.1; -
CC HSSP; P15116; INCT.
CC Genew; HGNC:8689; PCDH4.
CC MIM; 606330; -
CC MIM; 604967; -
CC InterPro: IPR002126; Cadherin.
CC Pfam: PF00028; cadherin; 5.
CC PRINTS; PR00205; CADHERIN.
CC SMART; SM00112; CA; 5.
CC PROSITE; PS00268; CADHERIN_1; 5.
CC PROSITE; PS00268; CADHERIN_2; 5.
CC Calcium-binding; Cell adhesion; Glycoprotein; Signal; Repeat;
CC Transmembrane; Multigene family.
CC SIGNAL 1 27
CC CHAIN 28 795
CC DOMAIN 28 689
CC TRANSMEM 690 710
CC DOMAIN 711 795
CC DOMAIN 34 132
CC DOMAIN 137 241
CC DOMAIN 246 346
CC DOMAIN 351 450
CC DOMAIN 455 560
CC DOMAIN 567 670
CC CARBOHYD 183 183
CC CARBOHYD 417 417
CC CARBOHYD 435 435
CC CARBOHYD 566 566
CC SEQUENCE 795 AA; 87269 MW; 32A521247DD6A56F CRC64;

Query Match 68.7%; Score 2205; DB 1; Length 795;
Best Local Similarity 69.5%; Pred. No. 1.3e-139;
Matches 428; Conservative 70; Mismatches 118; Indels 0; Gaps 0;

QY 12 ROYSLICVFLGVMAGAEPLRYAFVARETERGFLANLAIDLGIGVELSARCRIVSDPT 71
Db 11 ROYSLICVFLGVMAGAEPLRYAFVARETERGFLANLAIDLGIGVELSARCRIVSDPT 70
QY 72 IGFLLNPLTGDLINLEKIDREELCGPTEPCVYFOLLKEFQIRAEIWRVINDISP 131
Db 71 KORLQDRQTDGLLREKIDREELCGPTEPCVYFOLLKEFQIRAEIWRVINDISP 130
QY 132 VFIDREITLLESTPGATLLESABSDVGINLKNYITSSNYFEHINVDNGENVY 191
Db 131 IPEREVELIKLENSQPGFLPILIAEDVDVGSNGLOKTYTISPNSHFILRNHSEKKY 190
QY 192 SELVLDKVLDRREVEPELRLTLTGDSGSPRSCTTLRLIVLDINDNPFEVESLYKQV 251
Db 191 PDLVQKPLDRERQPEFSLTVALDGGSPRSCTVWVRLINDINDNAEFVHTPYGVQV 250
QY 252 PENSVPGLVTVVSARDLDTGSNGEIVAFYATERLTPRINSTGNLHKALENEA 311
Db 251 LENSPLDSPVRLARDIDAGNFGSVYGLFQASDEIKQTFEINVEYTGELILKKLDEK 310
QY 312 IQTYTLTQADGGGLSKCTVYVHVTDINDNPPELLMSSITSPENSPEVTVAVFRIR 371
Db 311 IKSYPEIATDGGGLSGKGVTVIEVDVNDNPPELLITSSISIPENAPETVVSIFRIR 370
QY 372 DBDSGNNAKWCYSIODHLPFLKPSVENFYTLVTERALDREREYNTITTYTDGTRRL 431
Db 371 DRDSGNNAKWCYSIODHLPFLKPSVENFYTLVTERALDREREYNTITTYTDGTRRL 430
QY 432 KTOHNLTVSVVNDNAPFESQTTTYTLRVRENSPALHIGSVASADRGSGANAOVTVSL 491
Db 431 KTOHNLTVSVVNDNAPFESQTTTYTLRVRENSPALHIGSVASADRGSGANAOVTVSL 490

```

QY 492 PPDPOLPLGSLVSNADNGOLFALSLDFEALQAFEFVGAADRCSPALSSOALVRYV 551  
 DB 491 PPDPPLPLASLVSINADNGHFLALSLDYEALQAFEFVGAADRCSPALSSOALVRYV 550  
 QY 552 ADANDNAPVLYPLONGSAPCTELVPRAAEAGYLVAKVAVDGDSCGNMYSYLKATE 611  
 DB 551 LDNDNSPEFLYPLONGSAPCTELVPRAAEAGYLVAKVAVDGDSCGNMYSYLKATE 610  
 QY 612 PGLFGVWANGEVRTA 627  
 DB 611 PGLFGVWANGEVRTA 626

RESULT 10  
 CDB5\_HUMAN STANDARD; PRT; 795 AA.

AC Q9Y5F1;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Protocadherin beta 12 precursor (PCDH-beta12).  
 GN PCDHB12.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99308636; PubMed=10380929;  
 RA Wu Q., Maniatis T.;  
 RT "A striking organization of a large family of human neural cadherin-  
 RT like cell adhesion genes."  
 RL Cell 97:779-790(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21223055; PubMed=11322959;  
 RA Vahne K., Kools P., Vanden Eynde E., van Roy F.;  
 RT "The human and murine protocadherin-beta one-exon gene families show  
 RT high evolutionary conservation, despite the difference in gene  
 RT number."  
 RL FEBS Lett. 495:120-125(2001).  
 CC -1- FUNCTION: POTENTIAL CALCIUM-DEPENDENT CELL-ADHESION PROTEIN.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: AF152491; AADA3752.1; -  
 DR EMBL: AF217746; AAK51614.1; -  
 DR HSSP: P15116; INCI.  
 DR Genew: HGNC:8683; PCDHB12.  
 DR MIM: 606338; -  
 DR MIM: 604967; -  
 DR InterPro: IPR002126; Cadherin.  
 DR Pfam: PF00028; cadherin.5.  
 DR PRINTS: PR00205; CADHERIN.  
 DR SMART: SM00112; CA; 6.  
 DR PROSITE: PS00232; CADHERIN.1; 5.  
 DR PROSITE: PS50268; CADHERIN.2; 6.  
 KW Calcium-binding; Cell adhesion; Glycoprotein; Signal; Repeat;  
 KW Transmembrane; Multigene family.  
 FT SIGNAL 1 26  
 FT CHAIN 27 795 POTENTIAL.  
 FT DOMAIN 27 690 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 691 711 POTENTIAL.  
 FT DOMAIN 712 795 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 35 133 CADHERIN 1.  
 FT DOMAIN 138 242 CADHERIN 2.  
 FT DOMAIN 247 347 CADHERIN 3.  
 FT DOMAIN 352 451 CADHERIN 4.  
 FT DOMAIN 456 561 CADHERIN 5.  
 FT DOMAIN 568 671 CADHERIN 6.  
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 436 436 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 795 AA; 86769 MW; 0A32163F81A0D37C CRC64;

Query Match: 68.0%; Score 2181; DB 1; Length 795;  
 Best Local Similarity: 68.7%; Pred. No. 5e-138;  
 Matches 431; Conservative 70; Mismatches 126; Indels 0; Gaps 0;

QY 1 MEARYVHALQKROYSLLCVFLGVSAGAEPLRYFAETERTGFLANLAIDLGLVEELS 60  
 DB 1 MENGAGTGLQIRVYLLFVLGMSQAGSETGNFLMEELQSGFVGNMLAKTLGLEVELS 60  
 QY 61 ARGCIVSDERTIGFLNPLPLTGDLILNKLDRBELCPTPCVLPQOLLEKPPQITRAE 120  
 DB 61 SRGARVVSNDNKECQLQDTNFTGDLILREMLDRBELCSNECVLYFVQLKNPQFLQIE 120  
 QY 121 LMYRDINDSHSPVFLDREITLNLTESTPGATFLLESADHSDVGINNRYNTISSNVFHI 180  
 DB 121 LQVRDINDSHSPVFLDREITLNLTESTPGATFLLESADHSDVGINNRYNTISSNVFHI 180  
 QY 181 NVHDNGEGNVSELYLDKVDREVPRLRLTLTGIDGSPRPSGTTILRLVLDINDNV 240  
 DB 181 KIRVPPDNRRKYPFLVLDKADYEERPELSFTLTALDGGSPRSGTALRVVVDINDNSP 240  
 QY 241 EFVESLKYVOYPPNSPVSLVTVASAPDLGSGNGELVYAFVYETRTLTFRINSNGN 300  
 DB 241 EFVEAFYEVKILENSIGSLVTVASAPDLGSGNGELVYAFVYETRTLTFRINSNGN 300  
 QY 301 LHLKAEVLAIGTYTLTIQAKDGGGLSGKCTVVVHTDINDNPBELMSLTPSPENS 360  
 DB 301 ITLAPLDFEALIESYTIQATDGGGLFGKSTVAYQWVDVNDAPELTVSSITSPRENT 360  
 QY 361 PEYVAVFRIRDRSGNNAKAVCSIOHLPFLVAKPSVENVYTLTERALDRERTENIT 420  
 DB 361 PEYVAVFRIRDRSGNNAKAVCSIOHLPFLVAKPSVENVYTLTERALDRERTENIT 420  
 QY 421 ITVVDLGPRLKTHNTITVSDVNDNAPPSQTTTLRVENNSPALHIGSVATPRDS 480  
 DB 421 ITVVDLGPRLKTHNTITVSDVNDNAPPSQTTTLRVENNSPALHIGSVATPRDS 480  
 QY 481 GANAQVTVSLPPDPOLPLGSLVSNADNGOLFALSLDFEALQAFEFVGAADRCSPA 540  
 DB 481 GTNAQVTVSLPPDPOLPLGSLVSNADNGOLFALSLDFEALQAFEFVGAADRCSPA 540  
 QY 541 LSSQALVAVLVADANDNAPVLYPLONGSAPCTELVPRAAEAGYLVAKVAVDGDSCQNA 600  
 DB 541 LSSQALVAVLVADANDNAPVLYPLONGSAPCTELVPRAAEAGYLVAKVAVDGDSCQNA 600  
 QY 601 WLSTQOLKATEPGLFGVWANGEVRTA 627  
 DB 601 WLSTQOLKATEPGLFGVWANGEVRTA 627

RESULT 11  
 CDB5\_HUMAN STANDARD; PRT; 795 AA.

AC Q9Y5E4; Q9YFV9;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Protocadherin beta 5 precursor (PCDH-beta5).  
 GN PCDHB5.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

	Matches 428;	Conservative	69;	Mismatches	129;	Indels	1;	Gaps
QY	1	MEARVHAHQKROVSLCYFLGVSNAGAEPLRYVAEETERTFLANLADLGLVEELS	60					
Db	1	METALAKTPQKROVMFLATLLIMFEGSEAVRISIPETESGVSANLKDGLGVGEIA	60					
QY	61	ARGCRIVSEETIGELLNPLTGDLLINKLDREELQEPTEPCVLPOLLERKFOIFRAE	12					
Db	61	TRGARHMYGKNGKELLQDLDTKTGNLLYBKLDRVWCATERPCHLPHRLLNVPQFPQD	12					
QY	121	LWVNDINDHSPVFLDREITLNLTESTPTGATFLLESADSDVGINLRNYLTSSNVEPHI	18					
Db	121	LQTLINDHAPERPEKEMULKIPESIQGVTEPKIAQDDPDIGSNVQNTTSPNSHPHV	18					
QY	181	NVNDNGEAGVNSVLYDKLVLDREYPERLRLTLTGDSGSPRRSGTLLIRLVLDINDVP	24					
Db	181	ATNRRDQDKRPELPYLDKRLDREERPELSLTLTLDGGAAPRSGGTTIRLVLDINDNAP	24					
QY	241	EFVESLYKQVPENSPVGSLVWTVYASRDLDTGSGEIVYAFYATERLTKTFRINSTGN	30					
Db	241	EFLOSTYEVQVPENSPLNSLVYASRDLGAGVAYALFQGDVET-QPEYIDEKTAE	29					
QY	301	LHLKAELETAIQYTLTIOAKDGGGSGKCTVYVHVTDINDNPPELLMSSLTSPENS	36					
Db	300	IRKRALDEEAQPPYVVEIVATDGGGSGSKCTVAIEVVDVNDNAPELTMTSLSPTEPA	35					
QY	361	PEVVAVFPIRRDGSNNAKWCSIQDHLPEVLCPSVENYTLVTERALDREREYNT	42					
Db	360	PEVVAVFVSVDPSDGNKMCISIQDHLPELPLKNTKNTLVLTQRTLDRESQAEYNT	41					
QY	421	ITVDTGTPRLKTOHNLVTVYSDVNDNAPRFSQTTTLRVRENNSPALHIGSVATDRDS	48					
Db	420	ITVDTGTPRLKTEHNITVLVSDVNDNAPAFQTSYTLFVRENNSPALHIGSVATDRDS	47					
QY	481	GAACQTVSELLRPHDPOPLSGISVSTANDNQFLFALSLDFEALQAFERRVGAADRGSA	54					
Db	480	GTNAQVTVSLPQNHRLKSLVSIINDNSHFLFALSLDYLEALQAFERRVGAATDRGSA	53					
QY	541	LSSQALVRLVADADNAPFVLYPLQNGSACTELVPRRAEAGYLVAKVAVADGSGQNA	60					
Db	540	LSSEALVRLVLDANDNSPFVLYPLQNGSACETELVPRRAEPGLVTKVAVADGSGQNA	59					
QY	601	WLSYQLLKATEPGLFGVAHNGEVRTA 627						
Db	600	WLSYQLLKATEPGLFSMAHNGEVRTA 626						
RESULT 12								
CDDB6_HUMAN		STANDARD:		PR1:		794	AA.	
AC	Q9YSE3;							
DT	16-OCT-2001 (Rel. 40, Created)							
DT	16-OCT-2001 (Rel. 40, Last sequence update)							
DT	15-JUN-2002 (Rel. 41, Last annotation update)							
DE	Protocadherin beta 6 precursor (PCDH-beta6).							
GN	PCDH6.							
OS	Homo sapiens (Human).							
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
OX	NCBI_TaxID=9606;							
RN	[1]							
RP	SEQUENCE FROM N.A.							
RP	MEDLINE=99308636; PubMed=10380929;							
RT	Wu Q., Maniatis T.,							
RT	A striking organization of a large family of human neural cadherin-							
RT	like cell adhesion genes.							
RT	Cell 97:779-790(1999).							
RP	SEQUENCE FROM N.A.							
RP	MEDLINE=21223055; PubMed=11332959;							
RT	Vanhalst K., Koois P., Vanden Eynde E., van Roy F.,							
RT	The human and murine protocadherin-beta one-exon gene families show							
RT	high evolutionary conservation, despite the difference in gene							



Query Match	67.88;	Score 2174.5;	DB 1;	Length 794;
Best Local Similarity	68.68;	Pred. No. 1.4e-137;		
Matches 430; Conservative	69;	Mismatches 125;	Indels 3;	Gaps 2

[illegible]

RESULT 13		
CDBB_HUMAN	STANDARD:	PRT: 797 AA.
ID CDBB_HUMAN		
AC Q9Y5F2;		
DT 16-OCT-2001 (Rel. 40, Created)		
DT 16-OCT-2001 (Rel. 40, Last sequence update)		
DT 15-JUN-2002 (Rel. 41, Last annotation update)		
DE Protocadherin beta 11 precursor (PCDH-beta11).		

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99308636; PubMed=10380929;
RA Wu Q., Maniatis T.;
RT "A striking organization of a large family of human neural cadherin-
RT like cell adhesion genes.";
RL Cell 97:779-790(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21223055; PubMed=11322959;
RA Vanhaelst K., Koels P., Vanden Eynde E., van Roy F.;
RT "The human and murine protocadherin-beta one-exon gene families show
RT high evolutionary conservation, despite the difference in gene
RT number.";
RL FEBS Lett. 495:120-125(2001).
CC -I- FUNCTION: POTENTIAL CALCIUM-DEPENDENT CELL-ADHESION PROTEIN.
CC -I- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -I- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF152490; AAD43751.1; -
DR EMBL: AF217747; AAK51615.1; -
DR HSSP: P15116; JNCJ
DR Genew; HGNC:8682; PCDHB11.
DR MIM: 606337; -
DR MIM: 604967; -
DR InterPro: IPR002126; Cadherin.
DR Pfam: PF00028; cadherin; 5.
DR PRINTS: PR00205; CADHERIN.
DR SMART: SM00112; CA; 6.
DR PROSITE: PS00232; CADHERIN_1; 5.
DR PROSITE: PS50268; CADHERIN_2; 6.
DR

```

KW Calcium-binding; Cell adhesion; Glycoprotein; Signal; Repeat;  
 Transmembrane; Multigene family.  
 FT SIGNAL 1 26 POTENTIAL.  
 FT CHAIN 27 797 PROTOCADHERIN BETA 11.  
 FT DOMAIN 27 690 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 691 711 POTENTIAL.  
 FT DOMAIN 712 797 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 35 133 CADHERIN 1.  
 FT DOMAIN 138 242 CADHERIN 2.  
 FT DOMAIN 247 347 CADHERIN 3.  
 FT DOMAIN 352 451 CADHERIN 4.  
 FT DOMAIN 456 561 CADHERIN 5.  
 FT DOMAIN 568 671 CADHERIN 6.  
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 436 436 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 797 AA; 87087 MW; 8FBADBC2858504 CRC64;

Query Match 67.7%; Score 2171; DB 1; Length 797;  
 Best Local Similarity 68.3%; Pred. No. 2.3e-137;  
 Matches 428; Conservative 76; Mismatches 123; Indels 0; Gaps 0;

QY 1 MEARYVHALOKROYSLLCVPLGVSWAGAPLRYFAEETERTFLANLADLGVEELS 60  
 DB 1 MENOCTPQOIRQVLLLVLLGMSQAGEIWSFVAEEMSGSEFVGNLAKDGLKVRRLS 60  
 QY 61 ARGGRVSDETIGLLPLNLTGDLINKELREELCGTEPCVLPFOELLEKQIFRAE 120  
 DB 61 SRGARVSNDRKKRQQLDININGDILLSETLDREELCGSIEPCVHLQVLMQNPFOLE 120  
 QY 121 LMRVDINHSVFLDRETTINLESTTPGATFLBSANDSVGNNLNRYTSSNVYFHI 180  
 DB 121 LQVRDINHSIFSEKQMLREIPENSIVGAVFLBSANDLVYGNVAVSTISPSHFHI 180  
 QY 181 NVHNGEGNVYSELVDVDRREYELRLTLTGDSGSPRSCTTLRIIYLOINDVNP 240  
 DB 181 KMRVYDPRKRPPELVLDALDYELPELSTLSALDGSPPSGTALRVVAVDINDNSP 240  
 QY 241 EFVESLKVQVPENSPVGLVVTVSARDIDGNGEIVYAFYATERTLKTFRINSTGN 300  
 DB 241 EFEQAFYEVKIRENSILIGSLIIVSAMOLDGTEGCEICTFSHSEDRLKFEINQKGE 300  
 QY 301 LHKAEINVEAIQYTTLTIOAKDGGGSGKCTVYVHTVDINDPELLMSITSPENS 360  
 DB 301 ITRAPLDEFIEISYIIIOATDGGGFGKSTVIIVHDVNDNAPELIVSSITSPENP 360  
 QY 361 PETVAAPRIRDRSGNNAKWCSTIODLPLVLPKSPVENFTLVTERALDEEREYNT 420  
 DB 361 PETVWVFSTIODISGNGRIVCSIPEDLPVLLSSVENYITLTERFLDRETAENYIT 420  
 QY 421 ITVTDLGPRLKTOHNTVTVSDVNDNAPTFSSQTYTLRVENNSPALHIGSVATDRS 480  
 DB 421 ITVTDLGPRLKTEHNTVTVSDVNDNAPTFQTSYTLFVRENNSPALHIGSVATDRS 480  
 QY 481 GANNOVYSLIPHPDPLGLSVISINDNGQLPRLSLDEALQAFERRGADRGSPA 540  
 DB 481 GTNNOVNSILIPDPLHPLASLVSINDNGHLPLRLSLDEALQAFERRGADRGSPA 540  
 QY 541 LSSQALVVLVADANDNAPFLYLPKONGSAPCTELVPRAAGYVAVVAVDGGSGNA 600  
 DB 541 LSSQALVVLVADANDNAPFLYLPKONGSAPCTELVPRAAGYVAVVAVDGGSGNA 600  
 QY 601 WLSTYOLKATEPGLGVVANGEVRTA 627  
 DB 601 WLSTYOLKATEPGLGVVANGEVRTA 627  
 RESULT 14  
 CDB8\_HUMAN STANDARD; PRT; 801 AA.  
 AC CDB8\_HUMAN  
 ID Q9UN66;  
 DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DT Protocadherin beta 8 precursor (PCDH-beta8) (Protocadherin 31).  
 GN PCDBB8 OR PCDH31.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_Taxid=9606;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99308636; PubMed=10380929;  
 RA Wu Q., Maniatis T.;  
 RT "A striking organization of a large family of human neural  
 RT cadherin-like cell adhesion genes";  
 RN Cell 97:779-790(1999).  
 RP REVISIONS.  
 RA Wu Q., Maniatis T.;  
 RN Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 RL [3]  
 RP SEQUENCE OF 192-801 FROM N.A.  
 RA Kools P.F.J., van Roy F.M.;  
 RT "Molecular analysis of the human protocadherin-3 (PCDH-beta) gene  
 RT cluster";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: POTENTIAL CALCIUM-DEPENDENT CELL-ADHESION PROTEIN.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF152501; AAD43762.2; -  
 CC EMBL; AF282973; AAG10031.1; -  
 CC Genew; HGNC:8693; PCDBB8.  
 CC MIM; 606334; -  
 CC MIM; 604967; -  
 DR InterPro: IPR002126; Cadherin.  
 DR Pfam; PF00028; cadherin. 5.  
 DR PRINTS; PR00205; CADHERIN.  
 DR SMART; SM00112; CA; 6.  
 DR PROSITE; PS00232; CADHERIN\_1; 5.  
 DR PROSITE; PS50268; CADHERIN\_2; 5.  
 KW Calcium-binding; Cell adhesion; Glycoprotein; Signal; Repeat;  
 KW Transmembrane; Multigene family.  
 FT SIGNAL 1 28 POTENTIAL.  
 FT CHAIN 29 801 PROTOCADHERIN BETA 8.  
 FT DOMAIN 29 691 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 692 710 POTENTIAL.  
 FT DOMAIN 711 801 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 36 134 CADHERIN 1.  
 FT DOMAIN 139 243 CADHERIN 2.  
 FT DOMAIN 248 348 CADHERIN 3.  
 FT DOMAIN 353 452 CADHERIN 4.  
 FT DOMAIN 457 562 CADHERIN 5.  
 FT DOMAIN 569 672 CADHERIN 6.  
 FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 568 568 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 199 199 N -> K (IN REF. 3).  
 FT CONFLICT 234 234 V -> L (IN REF. 3).  
 FT CONFLICT 244 244 Q -> E (IN REF. 3).  
 FT CONFLICT 244 244 G -> S (IN REF. 3).  
 FT CONFLICT 569 569 Q -> H (IN REF. 3).  
 FT CONFLICT 639 639 L -> V (IN REF. 3).  
 FT CONFLICT 661 661 F -> L (IN REF. 3).  
 FT CONFLICT 767 767  
 SO SEQUENCE 801 AA; 87623 MW; 30FB615CADA3A051 CRC64;



```

OY 360 SPETVAVFRINDRDSGNNAKWCSTIOHLPFVLKPSVENFTLVTERALDREERTENI 419
:|||||: | |||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||
Db 361 APEYVALFVSDDLDSGNGKISCSIOEDLPFLK-SAEFTYLLTERPLDRESRAEYNI 419
OY 420 TITVTDIGPRLKTOHNLVTVSDVNDNAPFTSOTTYTIRVRENNSPALHIGSVSATDRD 479
||||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 420 TITVTDIGPRLKTOHNLVTVSDVNDNAPFTSOTTYTIRVRENNSPALHIGSVSATDRD 479
OY 480 SGANQVYTSLLPDPOLPLGSLVSINADNGOLFALRSIDFEALQAFEFYRGADRGSP 539
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 480 SGTNQVYTSLLPDPOLPLGSLVSINADNGOLFALRSIDFEALQAFEFYRGADRGSP 539
OY 540 ALSSQALVRVAVDANDNAPFTVLPLONGSAPCTELVPRAEAGYLVAKVAVDGDGON 599
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 540 ALSSQALVRVAVDANDNAPFTVLPLONGSAPCTELVPRAEAGYLVAKVAVDGDGON 599
OY 600 AWLSTQLLKATEPGLGFWAHNGEVRTA 627
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 600 AWLSTQLLKATEPGLGFWAHNGEVRTA 627

```

Search completed: May 16, 2003, 10:46:24  
 Job time : 19 secs